THIS PAGE BLANK (USPTO)

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/47, C12N 15/10, 15/66, C12Q 1/68, G01N 33/50, C07K 16/18, G01N 33/53, A61K 48/00, 38/17

A2

(11) International Publication Number: WO 99/06551

(43) International Publication Date: 11 February 1999 (11.02.99)

(21) International Application Number: PCT/IB98/01235

(22) International Filing Date: 31 July 1998 (31.07.98)

(30) Priority Data:
08/905,133 1 August 1997 (01.08.97) US

(71) Applicant (*for all designated States except US*): GENSET [FR/FR]; 24, rue Royale, F-75008 Paris (FR).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): DUMAS MILNE EDWARDS, Jean-Baptiste [FR/FR]; 8, rue Grégoire de Tours, F-75006 Paris (FR). DUCLERT, Aymeric [FR/FR]; 6 ter, rue Victorine, F-94100 Saint-Maur (FR). LACROIX, Bruno [FR/FR]; 93, route de Vourles, F-69230 Saint-Genis Laval (FR).

(74) Agents: MARTIN, Jean-Jacques et al.; Cabinet Regimbeau, 26, avenue Kléber, F-75116 Paris (FR).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: 5' ESTs FOR SECRETED PROTEINS IDENTIFIED FROM BRAIN TISSUES

(57) Abstract

The sequence of 5' ESTs derived from mRNAs encoding secreted proteins are disclosed. The 5' ESTs may be to obtain cDNAS, and genomic DNAs corresponding to the 5' ESTs. The 5' ESTs may also be used in diagnostic, forensic, gene therapy, and chromosome mapping procedures. Upstream regulatory sequences may also be obtained using the 5' ESTs. The 5' ESTs may also be used to design expression vectors and secretion vectors.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

5' ESTs FOR SECRETED PROTEINS IDENTIFIED FROM BRAIN TISSUES

Background of the Invention

The estimated 50,000-100,000 genes scattered along the human chromosomes offer tremendous promise for the understanding, diagnosis, and treatment of human diseases. In addition, probes capable of specifically hybridizing to loci distributed throughout the human genome find applications in the construction of high resolution chromosome maps and in the identification of individuals.

In the past, the characterization of even a single human gene was a painstaking process, requiring years of effort. Recent developments in the areas of cloning vectors, DNA sequencing, and computer technology have merged to greatly accelerate the rate at which human genes can be isolated, sequenced, mapped, and characterized. Cloning vectors such as yeast artificial chromosomes (YACs) and bacterial artificial chromosomes (BACs) are able to accept DNA inserts ranging from 300 to 1000 kilobases (kb) or 100-400 kb in length respectively, thereby facilitating the manipulation and ordering of DNA sequences distributed over great distances on the human chromosomes. Automated DNA sequencing machines permit the rapid sequencing of human genes. Bioinformatics software enables the comparison of nucleic acid and protein sequences, thereby assisting in the characterization of human gene products.

Currently, two different approaches are being pursued for identifying and characterizing the genes distributed along the human genome. In one approach, large fragments of genomic DNA are isolated, cloned, and sequenced. Potential open reading frames in these genomic sequences are identified using bioinformatics software. However, this approach entails sequencing large stretches of human DNA which do not encode proteins in order to find the protein encoding sequences scattered throughout the genome. In addition to requiring extensive sequencing, the bioinformatics software may mischaracterize the genomic sequences obtained. Thus, the software may produce false positives in which non-coding DNA is mischaracterized as coding DNA or false negatives in which coding DNA is mislabeled as non-coding DNA.

An alternative approach takes a more direct route to identifying and characterizing human genes. In this approach, complementary DNAs (cDNAs) are synthesized from isolated messenger RNAs (mRNAs) which encode human proteins. Using this approach,

sequencing is only performed on DNA which is derived from protein coding portions of the genome. Often, only short stretches of the cDNAs are sequenced to obtain sequences called expressed sequence tags (ESTs). The ESTs may then be used to isolate or purify extended cDNAs which include sequences adjacent to the EST sequences. The extended cDNAs may contain all of the sequence of the EST which was used to obtain them or only a portion of the sequence of the EST which was used to obtain them. In addition, the extended cDNAs may contain the full coding sequence of the gene from which the EST was derived or, alternatively, the extended cDNAs may include portions of the coding sequence of the gene from which the EST was derived. It will be appreciated that there may be several extended cDNAs which include the EST sequence as a result of alternate splicing or the activity of alternative promoters.

In the past, these short EST sequences were often obtained from oligo-dT primed cDNA libraries. Accordingly, they mainly corresponded to the 3′ untranslated region of the mRNA. In part, the prevalence of EST sequences derived from the 3′ end of the mRNA is a result of the fact that typical techniques for obtaining cDNAs are not well suited for isolating cDNA sequences derived from the 5′ ends of mRNAs. (Adams *et al., Nature* 377:3-174, 1996; Hillier *et al., Genome Res.* 6:807-828, 1996).

In addition, in those reported instances where longer cDNA sequences have been obtained, the reported sequences typically correspond to coding sequences and do not include the full 5′ untranslated region of the mRNA from which the cDNA is derived. Such incomplete sequences may not include the first exon of the mRNA, particularly in situations where the first exon is short. Furthermore, they may not include some exons, often short ones, which are located upstream of splicing sites. Thus, there is a need to obtain sequences derived from the 5′ ends of mRNAs.

While many sequences derived from human chromosomes have practical applications, approaches based on the identification and characterization of those chromosomal sequences which encode a protein product are particularly relevant to diagnostic and therapeutic uses. Of the 50,000-100,000 protein coding genes, those genes encoding proteins which are secreted from the cell in which they are synthesized, as well as the secreted proteins themselves, are particularly valuable as potential therapeutic agents. Such proteins are often

involved in cell to cell communication and may be responsible for producing a clinically relevant response in their target cells.

In fact, several secretory proteins, including tissue plasminogen activator, G-CSF, GM-CSF, erythropoietin, human growth hormone, insulin, interferon-α, interferon-β, interferon-γ, and interleukin-2, are currently in clinical use. These proteins are used to treat a wide range of conditions, including acute myocardial infarction, acute ischemic stroke, anemia, diabetes, growth hormone deficiency, hepatitis, kidney carcinoma, chemotherapy induced neutropenia and multiple sclerosis. For these reasons, extended cDNAs encoding secreted proteins or portions thereof represent a particularly valuable source of therapeutic agents. Thus, there is a need for the identification and characterization of secreted proteins and the nucleic acids encoding them.

In addition to being therapeutically useful themselves, secretory proteins include short peptides, called signal peptides, at their amino termini which direct their secretion. These signal peptides are encoded by the signal sequences located at the 5′ ends of the coding sequences of genes encoding secreted proteins. Because these signal peptides will direct the extracellular secretion of any protein to which they are operably linked, the signal sequences may be exploited to direct the efficient secretion of any protein by operably linking the signal sequences to a gene encoding the protein for which secretion is desired. In addition, portions of signal sequences may also be used to direct the intracellular import of a peptide or protein of interest. This may prove beneficial in gene therapy strategies in which it is desired to deliver a particular gene product to cells other than the cell in which it is produced. Signal sequences encoding signal peptides also find application in simplifying protein purification techniques. In such applications, the extracellular secretion of the desired protein greatly facilitates purification by reducing the number of undesired proteins from which the desired protein must be selected. Thus, there exists a need to identify and characterize the 5′ portions of the genes for secretory proteins which encode signal peptides.

Public information on the number of human genes for which the promoters and upstream regulatory regions have been identified and characterized is quite limited. In part, this may be due to the difficulty of isolating such regulatory sequences. Upstream regulatory sequences such as transcription factor binding sites are typically too short to be utilized as probes for isolating promoters from human genomic libraries. Recently, some approaches

have been developed to isolate human promoters. One of them consists of making a CpG island library (Cross, *et al., Nature Genetics* 6: 236-244, 1994). The second consists of isolating human genomic DNA sequences containing SpeI binding sites by the use of SpeI binding protein. (Mortlock *et al., Genome Res.* 6:327-335, 1996). Both of these approaches have their limits due to a lack of specificity or of comprehensiveness.

The present 5′ ESTs may be used to efficiently identify and isolate upstream regulatory regions which control the location, developmental stage, rate, and quantity of protein synthesis, as well as the stability of the mRNA. (Theil, *BioFactors* 4:87-93, 1993). Once identified and characterized, these regulatory regions may be utilized in gene therapy or protein purification schemes to obtain the desired amount and locations of protein synthesis or to inhibit, reduce, or prevent the synthesis of undesirable gene products.

In addition, ESTs containing the 5′ ends of secretory protein genes may include sequences useful as probes for chromosome mapping and the identification of individuals. Thus, there is a need to identify and characterize the sequences upstream of the 5′ coding sequences of genes encoding secretory proteins.

Summary of the Invention

The present invention relates to purified, isolated, or recombinant ESTs which include sequences derived from the authentic 5′ ends of their corresponding mRNAs. The term "corresponding mRNA" refers to the mRNA which was the template for the cDNA synthesis which produced the 5′ EST. These sequences will be referred to hereinafter as "5′ ESTs." As used herein, the term "purified" does not require absolute purity; rather, it is intended as a relative definition. Individual 5′ EST clones isolated from a cDNA library have been conventionally purified to electrophoretic homogeneity. The sequences obtained from these clones could not be obtained directly either from the library or from total human DNA. The cDNA clones are not naturally occurring as such, but rather are obtained via manipulation of a partially purified naturally occurring substance (messenger RNA). The conversion of mRNA into a cDNA library involves the creation of a synthetic substance (cDNA) and pure individual cDNA clones can be isolated from the synthetic library by clonal selection. Thus, creating a cDNA library from messenger RNA and subsequently isolating individual clones from that library results in an approximately 10^4-10^6 fold purification of the native message.

Purification of starting material or natural material to at least one order of magnitude, preferably two or three orders, and more preferably four or five orders of magnitude is expressly contemplated.

As used herein, the term "isolated" requires that the material be removed from its original environment (e.g., the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide present in a living animal is not isolated, but the same polynucleotide, separated from some or all of the coexisting materials in the natural system, is isolated.

As used herein, the term "recombinant" means that the 5′ EST is adjacent to "backbone" nucleic acid to which it is not adjacent in its natural environment. Additionally, to be "enriched" the 5′ ESTs will represent 5% or more of the number of nucleic acid inserts in a population of nucleic acid backbone molecules. Backbone molecules according to the present invention include nucleic acids such as expression vectors, self-replicating nucleic acids, viruses, integrating nucleic acids, and other vectors or nucleic acids used to maintain or manipulate a nucleic acid insert of interest. Preferably, the enriched 5′ ESTs represent 15% or more of the number of nucleic acid inserts in the population of recombinant backbone molecules. More preferably, the enriched 5′ ESTs represent 50% or more of the number of nucleic acid inserts in the population of recombinant backbone molecules. In a highly preferred embodiment, the enriched 5′ ESTs represent 90% or more of the number of nucleic acid inserts in the population of recombinant backbone molecules.

"Stringent", moderate," and "low" hybridization conditions are as defined in Example 29.

Unless otherwise indicated, a "complementary" sequence is fully complementary.

Thus, 5′ ESTs in cDNA libraries in which one or more 5′ ESTs make up 5% or more of the number of nucleic acid inserts in the backbone molecules are "enriched recombinant 5′ ESTs" as defined herein. Likewise, 5′ ESTs in a population of plasmids in which one or more 5′ EST of the present invention have been inserted such that they represent 5% or more of the number of inserts in the plasmid backbone are " enriched recombinant 5′ ESTs" as defined herein. However, 5′ ESTs in cDNA libraries in which 5′ ESTs constitute less than 5% of the number of nucleic acid inserts in the population of backbone molecules, such as libraries in

which backbone molecules having a 5′ EST insert are extremely rare, are not "enriched recombinant 5′ ESTs."

In particular, the present invention relates to 5′ ESTs which are derived from genes encoding secreted proteins. As used herein, a "secreted" protein is one which, when expressed in a suitable host cell, is transported across or through a membrane, including transport as a result of signal peptides in its amino acid sequence. "Secreted" proteins include without limitation proteins secreted wholly (e.g. soluble proteins), or partially (e.g. receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins which are transported across the membrane of the endoplasmic reticulum.

Such 5′ ESTs include nucleic acid sequences, called signal sequences, which encode signal peptides which direct the extracellular secretion of the proteins encoded by the genes from which the 5′ ESTs are derived. Generally, the signal peptides are located at the amino termini of secreted proteins.

Secreted proteins are translated by ribosomes associated with the "rough" endoplasmic reticulum. Generally, secreted proteins are co-translationally transferred to the membrane of the endoplasmic reticulum. Association of the ribosome with the endoplasmic reticulum during translation of secreted proteins is mediated by the signal peptide. The signal peptide is typically cleaved following its co-translational entry into the endoplasmic reticulum. After delivery to the endoplasmic reticulum, secreted proteins may proceed through the Golgi apparatus. In the Golgi apparatus, the proteins may undergo post-translational modification before entering secretory vesicles which transport them across the cell membrane.

The 5′ ESTs of the present invention have several important applications. For example, they may be used to obtain and express cDNA clones which include the full protein coding sequences of the corresponding gene products, including the authentic translation start sites derived from the 5′ ends of the coding sequences of the mRNAs from which the 5′ ESTs are derived. These cDNAs will be referred to hereinafter as "full length cDNAs." These cDNAs may also include DNA derived from mRNA sequences upstream of the translation start site. The full length cDNA sequences may be used to express the proteins corresponding to the 5′ ESTs. As discussed above, secreted proteins are therapeutically important. Thus, the proteins expressed from the cDNAs may be useful in treating or

controlling a variety of human conditions. The 5′ ESTs may also be used to obtain the corresponding genomic DNA. The term "corresponding genomic DNA" refers to the genomic DNA which encodes the mRNA from which the 5′ EST was derived.

Alternatively, the 5′ ESTs may be used to obtain and express extended cDNAs encoding portions of the secreted protein. The portions may comprise the signal peptides of the secreted proteins or the mature proteins generated when the signal peptide is cleaved off. The portions may also comprise polypeptides having at least 10 consecutive amino acids encoded by the extended cDNAs or full length cDNAs. Alternatively, the portions may comprise at least 15 consecutive amino acids encoded by the extended cDNAs or full length cDNAs. In some embodiments, the portions may comprise at least 25 consecutive amino acids encoded by the extended cDNAs or full length cDNAs. In other embodiments, the portions may comprise at least 40 amino acids encoded by the extended cDNAs or full length cDNAs.

Antibodies which specifically recognize the entire secreted proteins encoded by the extended cDNAs, full length cDNAs, or fragments thereof having at least 10 consecutive amino acids, at least 15 consecutive amino acids, at least 25 consecutive amino acids, or at least 40 consecutive amino acids may also be obtained as described below. Antibodies which specifically recognize the mature protein generated when the signal peptide is cleaved may also be obtained as described below. Similarly, antibodies which specifically recognize the signal peptides encoded by the extended cDNAs or full length cDNAs may also be obtained.

In some embodiments, the extended cDNAs obtained using the 5′ ESTs include the signal sequence. In other embodiments, the extended cDNAs obtained using the 5′ ESTs may include the full coding sequence for the mature protein (*i.e.* the protein generated when the signal polypeptide is cleaved off). In addition, the extended cDNAs obtained using the 5′ ESTs may include regulatory regions upstream of the translation start site or downstream of the stop codon which control the amount, location, or developmental stage of gene expression.

As discussed above, secreted proteins are therapeutically important. Thus, the proteins expressed from the extended cDNAs or full length cDNAs obtained using the 5′ ESTs may be useful in treating or controlling a variety of human conditions.

The 5′ ESTs (or cDNAs or genomic DNAs obtained therefrom) may be used in forensic procedures to identify individuals or in diagnostic procedures to identify individuals having genetic diseases resulting from abnormal expression of the genes corresponding to the 5′ ESTs. In addition, the present invention is useful for constructing a high resolution map of the human chromosomes.

The present invention also relates to secretion vectors capable of directing the secretion of a protein of interest. Such vectors may be used in gene therapy strategies in which it is desired to produce a gene product in one cell which is to be delivered to another location in the body. Secretion vectors may also facilitate the purification of desired proteins.

The present invention also relates to expression vectors capable of directing the expression of an inserted gene in a desired spatial or temporal manner or at a desired level. Such vectors may include sequences upstream of the 5′ ESTs, such as promoters or upstream regulatory sequences.

Finally, the present invention may also be used for gene therapy to control or treat genetic diseases. Signal peptides may also be fused to heterologous proteins to direct their extracellular secretion.

Bacterial clones containing Bluescript plasmids having inserts containing the 5′ ESTs of the present invention (SEQ ID NOs: 38-195 are presently stored at 80°C in 4% (v/v) glycerol in the inventor's laboratories under the designations listed next to the SEQ ID NOs in II). The inserts may be recovered from the deposited materials by growing the appropriate clones on a suitable medium. The Bluescript DNA can then be isolated using plasmid isolation procedures familiar to those skilled in the art such as alkaline lysis minipreps or large scale alkaline lysis plasmid isolation procedures. If desired the plasmid DNA may be further enriched by centrifugation on a cesium chloride gradient, size exclusion chromatography, or anion exchange chromatography. The plasmid DNA obtained using these procedures may then be manipulated using standard cloning techniques familiar to those skilled in the art. Alternatively, a PCR can be done with primers designed at both ends of the EST insertion. The PCR product which corresponds to the 5′ EST can then be manipulated using standard cloning techniques familiar to those skilled in the art.

One aspect of the present invention is a purified or isolated nucleic acid having the sequence of one of SEQ ID NOs: 38-195 or having a sequence complementary thereto. In one embodiment, the nucleic acid is recombinant.

Another aspect of the present invention is a purified or isolated nucleic acid comprising at least 10 consecutive bases of the sequence of one of SEQ ID NOs: 38-195 or one of the sequences complementary thereto.

Yet another aspect of the present invention is a purified or isolated nucleic acid comprising at least 15 consecutive bases of one of the sequences of SEQ ID NOs: 38-195 or one of the sequences complementary thereto. In one embodiment, the nucleic acid is recombinant.

A further aspect of the present invention is a purified or isolated nucleic acid of at least 15 bases capable of hybridizing under stringent conditions to the sequence of one of SEQ ID NOs: 38-195 or one of the sequences complementary to the sequences of SEQ ID NOs: 38-195. In one embodiment, the nucleic acid is recombinant.

Another aspect of the present invention is a purified or isolated nucleic acid encoding a human gene product, said human gene product having a sequence partially encoded by one of the sequences of SEQ ID NO: 38-195.

Still another aspect of the present invention is a method of making a cDNA encoding a human secretory protein, said human secretory protein being partially encoded by one of SEQ ID NOs 38-195, comprising the steps of contacting a collection of mRNA molecules from human cells with a primer comprising at least 15 consecutive nucleotides of a sequence complementary to one of SEQ ID NOs: 38-195; hybridizing said primer to an mRNA in said collection that encodes said protein; reverse transcribing said hybridized primer to make a first cDNA strand from said mRNA; making a second cDNA strand complementary to said first cDNA strand; and isolating the resulting cDNA encoding said protein comprising said first cDNA strand and said second cDNA strand.

Another aspect of the invention is an isolated or purified cDNA encoding a human secretory protein, said human secretory protein comprising the protein encoded by one of SEQ ID NOs 38-195 or a fragment thereof of at least 10 amino acids, said cDNA being obtainable by the method described in the preceding paragraph. In one embodiment, the

cDNA comprises the full protein coding sequence of said protein which sequence is partially included in one of the sequences of SEQ ID NOs: 38-195.

Another aspect of the present invention is a method of making a cDNA encoding a human secretory protein that is partially encoded by one of SEQ ID NOs 38-195, comprising the steps of obtaining a cDNA comprising one of the sequences of SEQ ID NOs: 38-195; contacting said cDNA with a detectable probe comprising at least 15 consecutive nucleotides of said sequence of SEQ ID NO: 38-195 or a sequence complementary thereto under conditions which permit said probe to hybridize to said cDNA; identifying a cDNA which hybridizes to said detectable probe; and isolating said cDNA which hybridizes to said probe.

Another aspect of the present invention is an isolated or purified cDNA encoding a human secretory protein, said human secretory protein comprising the protein encoded by one of SEQ ID NOs 38-195 or a fragment thereof of at least 10 amino acids, said cDNA being obtainable by the method described in the preceding paragraph. In one embodiment, the cDNA comprises the full protein coding sequence partially included in one of the sequences of SEQ ID NOs: 38-195.

Another aspect of the present invention is a method of making a cDNA comprising one of the sequence of SEQ ID NOs: 38-195, comprising the steps of contacting a collection of mRNA molecules from human cells with a first primer capable of hybridizing to the polyA tail of said mRNA; hybridizing said first primer to said polyA tail; reverse transcribing said mRNA to make a first cDNA strand; making a second cDNA strand complementary to said first cDNA strand using at least one primer comprising at least 15 nucleotides of one of the sequences of SEQ ID NOs 38-195; and isolating the resulting cDNA comprising said first cDNA strand and said second cDNA strand.

Another aspect of the present invention is an isolated or purified cDNA encoding a human secretory protein, said human secretory protein comprising the protein encoded by one of SEQ ID NOs 38-195 or a fragment thereof of at least 10 amino acids, said cDNA being obtainable by the method described in the preceding paragraph. In one embodiment, the cDNA comprises the full protein coding sequence partially included in one of the sequences of SEQ ID NOs: 38-195.

In one embodiment of the method described in the two paragraphs above, the second cDNA strand is made by contacting said first cDNA strand with a first pair of primers, said

first pair of primers comprising a second primer comprising at least 15 consecutive nucleotides of one of the sequences of SEQ ID NOs 38-195 and a third primer having a sequence therein which is included within the sequence of said first primer; performing a first polymerase chain reaction with said first pair of nested primers to generate a first PCR product; contacting said first PCR product with a second pair of primers, said second pair of primers comprising a fourth primer, said fourth primer comprising at least 15 consecutive nucleotides of said sequence of one of SEQ ID NOs: 38-195 , and a fifth primer, said fourth and fifth primers being capable of hybridizing to sequences within said first PCR product; and performing a second polymerase chain reaction, thereby generating a second PCR product.

One aspect of the present invention is an isolated or purified cDNA encoding a human secretory protein, said human secretory protein comprising the protein encoded by one of SEQ ID NOs 38-195, or a fragment thereof of at least 10 amino acids, said cDNA being obtainable by the method of the preceding paragraph. In one embodiment, the cDNA comprises the full protein coding sequence partially included in one of the sequences of SEQ ID NOs: 38-195.

Another aspect of the present invention is the method described four paragraphs above in which the second cDNA strand is made by contacting said first cDNA strand with a second primer comprising at least 15 consecutive nucleotides of the sequences of SEQ ID NOs: 38-195; hybridizing said second primer to said first strand cDNA; and extending said hybridized second primer to generate said second cDNA strand.

Another aspect of the present invention is an isolated or purified cDNA encoding a human secretory protein, said human secretory protein comprising the protein partially encoded by one of SEQ ID NOs 38-195 or comprising a fragment thereof of at least 10 amino acids, said cDNA being obtainable by the method described in the preceding paragraph. In one embodiment, the cDNA comprises the full protein coding sequence partially included in of one of the sequences of SEQ ID NOs: 38-195.

Another aspect of the present invention is a method of making a protein comprising one of the sequences of SEQ ID NOs: 196-353, comprising the steps of obtaining a cDNA encoding the full protein sequence partially included in one of the sequences of sequence of SEQ ID NOs: 38-195; inserting said cDNA in an expression vector such that said cDNA is

operably linked to a promoter; introducing said expression vector into a host cell whereby said host cell produces the protein encoded by said cDNA; and isolating said protein.

Another aspect of the present invention is an isolated protein obtainable by the method described in the preceding paragraph.

Another aspect of the present invention is a method of obtaining a promoter DNA comprising the steps of obtaining DNAs located upstream of the nucleic acids of SEQ ID NOs: 38-195 or the sequences complementary thereto; screening said upstream DNAs to identify a promoter capable of directing transcription initiation; and isolating said DNA comprising said identified promoter. In one embodiment, the obtaining step comprises chromosome walking from said nucleic acids of SEQ ID NOs: 38-195 or sequences complementary thereto. In another embodiment, the screening step comprises inserting said upstream sequences into a promoter reporter vector. In another embodiment, the screening step comprises identifying motifs in said upstream DNAs which are transcription factor binding sites or transcription start sites.

Another aspect of the present invention is an isolated promoter obtainable by the method described above.

Another aspect of the present invention is an isolated or purified protein comprising one of the sequences of SEQ ID NOs: 196-353.

Another aspect of the present invention is the inclusion of at least one of the sequences of SEQ ID NOs: 38-195, or one of the sequences complementary to the sequences of SEQ ID NOs: 38-195, or a fragment thereof of at least 15 consecutive nucleotides in an array of discrete ESTs or fragments thereof of at least 15 nucleotides in length. In one embodiment, the array includes at least two of the sequences of SEQ ID NOs: 38-195, the sequences complementary to the sequences of SEQ ID NOs: 38-195, or fragments thereof of at least 15 consecutive nucleotides. In another embodiment, the array includes at least five of the sequences of SEQ ID NOs: 38-195, the sequences complementary to the sequences of SEQ ID NOs: 38-195, or fragments thereof of at least 15 consecutive nucleotides.

Another aspect of the present invention is a promoter having a sequence selected from the group consisting of SEQ ID NOs: 31, 34, and 37.

Brief Description of the Drawings

Figure 1 is a summary of a procedure for obtaining cDNAs which have been selected to include the 5′ ends of the mRNAs from which they derived.

Figure 2 shows the distribution of Von Heijne scores for 5′ ESTs in each of the categories described herein and the probability that these 5′ ESTs encode a signal peptide.

Figure 3 summarizes a general method used to clone and sequence extended cDNAs containing sequences adjacent to 5' ESTs.

Figure 4 (description of promoters structure isolated from SignalTag 5' ESTs) provides a schematic description of promoters isolated and the way they are assembled with the corresponding 5′ tags.

Detailed Description of the Preferred Embodiment

Table IV is an analysis of the 43 amino acids located at the N terminus of all human SwissProt proteins to determine the frequency of false positives and false negatives using the techniques for signal peptide identification described herein.

Table V shows the distribution of 5′ ESTs in each category described herein and the number of 5′ ESTs in each category having a given minimum Von Heijne's score.

Table VI shows the distribution of 5′ ESTs in each category described herein with respect to the tissue from which the 5' ESTs of the corresponding mRNA were obtained.

Table VII describes the transcription factor binding sites present in each of these promoters.

I. General Methods for Obtaining 5′ ESTs derived from mRNAs with intact 5' ends

In order to obtain the 5' ESTs of the present invention, mRNAs with intact 5' ends must be obtained. Currently, there are two approaches for obtaining such mRNAs with intact 5' ends as described below: either chemical (1) or enzymatic (2).

1. Chemical Methods for Obtaining mRNAs having Intact 5′ Ends

One of these approaches is a chemical modification method involving derivatization of the 5′ ends of the mRNAs and selection of the derivatized mRNAs. The 5′ ends of eukaryotic mRNAs possess a structure referred to as a "cap" which comprises a guanosine

methylated at the 7 position. The cap is joined to the first transcribed base of the mRNA by a 5′, 5′-triphosphate bond. In some instances, the 5′ guanosine is methylated in both the 2 and 7 positions. Rarely, the 5′ guanosine is trimethylated at the 2, 7 and 7 positions. In the chemical method for obtaining mRNAs having intact 5′ ends, the 5′ cap is specifically derivatized and coupled to a reactive group on an immobilizing substrate. This specific derivatization is based on the fact that only the ribose linked to the methylated guanosine at the 5′ end of the mRNA and the ribose linked to the base at the 3′ terminus of the mRNA, possess 2′, 3′-cis diols.

Optionally, the 2′, 3′-cis diol of the 3′ terminal ribose may be chemically modified, substituted, converted, or eliminated, leaving only the ribose linked to the methylated guanosine at the 5′ end of the mRNA with a 2′, 3′-cis diol. A variety of techniques are available for eliminating the 2′, 3′-cis diol on the 3′ terminal ribose. For example, controlled alkaline hydrolysis may be used to generate mRNA fragments in which the 3′ terminal ribose is a 3′-phosphate, 2′-phosphate or (2′, 3′)-cyclophosphate. Thereafter, the fragment which includes the original 3′ ribose may be eliminated from the mixture through chromatography on an oligodT column. Alternatively, a base which lacks the 2′, 3′-cis diol may be added to the 3′ end of the mRNA using an RNA ligase such as T4 RNA ligase. Example 1 below describes a method for ligation of a nucleoside diphosphate to the 3′ end of messenger RNA.

EXAMPLE 1

Ligation of the Nucleoside Diphosphate pCp to the 3' End of mRNA.

One μg of RNA was incubated in a final reaction medium of 10 μl in the presence of 5 U of T_4 phage RNA ligase in the buffer provided by the manufacturer (Gibco - BRL), 40 U of the RNase inhibitor RNasin (Promega) and, 2 μl of ^{32}pCp (Amersham #PB 10208). The incubation was performed at 37°C for 2 hours or overnight at 7-8°C.

Following modification or elimination of the 2′, 3′-cis diol at the 3′ ribose, the 2′, 3′-cis diol present at the 5′ end of the mRNA may be oxidized using reagents such as $NaBH_4$, $NaBH_3CN$, or sodium periodate, thereby converting the 2′, 3′-cis diol to a dialdehyde.

Example 2 describes the oxidation of the 2′, 3′-cis diol at the 5′ end of the mRNA with sodium periodate.

EXAMPLE 2

Oxidation of 2′, 3′-cis diol at the 5′ End of the mRNA with Sodium Periodate

0.1 OD unit of either a capped oligoribonucleotide of 47 nucleotides (including the cap) or an uncapped oligoribonucleotide of 46 nucleotides were treated as follows. The oligoribonucleotides were produced by *in vitro* transcription using the transcription kit "AmpliScribe T7" (Epicentre Technologies). As indicated below, the DNA template for the RNA transcript contained a single cytosine. To synthesize the uncapped RNA, all four NTPs were included in the *in vitro* transcription reaction. To obtain the capped RNA, GTP was replaced by an analogue of the cap, m7G(5′)ppp(5′)G. This compound, recognized by the polymerase, was incorporated into the 5′ end of the nascent transcript during the initiation of transcription but was not incorporated during the extension step. Consequently, the resulting RNA contained a cap at its 5′ end. The sequences of the oligoribonucleotides produced by the *in vitro* transcription reaction were:

+Cap:

5′m7GpppGCAUCCUACUCCCAUCCAAUUCCACCCUAACUCCUCCCAUCUCCAC-3′ (SEQ ID NO:1)

-Cap:

5′-pppGCAUCCUACUCCCAUCCAAUUCCACCCUAACUCCUCCCAUCUCCAC-3′ (SEQ ID NO:2)

The oligoribonucleotides were dissolved in 9 µl of acetate buffer (0.1 M sodium acetate, pH 5.2) and 3 µl of freshly prepared 0.1 M sodium periodate solution. The mixture was incubated for 1 hour in the dark at 4°C or room temperature. Thereafter, the reaction was stopped by adding 4 µl of 10% ethylene glycol. The product was ethanol precipitated, resuspended in at least 10 µl of water or appropriate buffer and dialyzed against water.

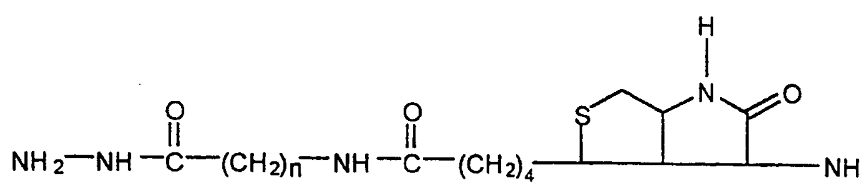
The resulting aldehyde groups may then be coupled to molecules having a reactive amine group, such as hydrazine, carbazide, thiocarbazide or semicarbazide groups, in order to facilitate enrichment of the 5′ ends of the mRNAs. Molecules having reactive amine groups

which are suitable for use in selecting mRNAs having intact 5′ ends include avidin, proteins, antibodies, vitamins, ligands capable of specifically binding to receptor molecules, or oligonucleotides. Example 3 below describes the coupling of the resulting dialdehyde to biotin.

EXAMPLE 3

Coupling of the Dialdehyde at the 5' End of Transcripts with Biotin

The oxidation product obtained in Example 2 was dissolved in 50 µl of sodium acetate at a pH between 5 and 5.2 and 50 µl of freshly prepared 0.02 M solution of biotin hydrazide in a methoxyethanol/water mixture (1:1) of formula:



In the compound used in these experiments, n=5. However, it will be appreciated that other commercially available hydrazides may also be used, such as molecules of the above formula in which n varies from 0 to 5. The mixture was then incubated for 2 hours at 37°C, precipitated with ethanol and dialyzed against distilled water. Example 4 demonstrates the specificity of the biotinylation reaction.

EXAMPLE 4

Specificity of Biotinylation of Capped Transcripts

The specificity of the biotinylation for capped mRNAs was evaluated by gel electrophoresis of the following samples:

Sample 1. The 46 nucleotide uncapped *in vitro* transcript prepared as in Example 2 and labeled with ^{32}pCp as described in Example 1.

Sample 2. The 46 nucleotide uncapped *in vitro* transcript prepared as in Example 2, labeled with ^{32}pCp as described in Example 1, treated with the oxidation reaction of Example 2, and subjected to the biotinylation conditions of Example 3.

Sample 3. The 47 nucleotide capped *in vitro* transcript prepared as in Example 2 and labeled with ^{32}pCp as described in Example 1.

Sample 4. The 47 nucleotide capped *in vitro* transcript prepared as in Example 2, labeled with ^{32}pCp as described in Example 1, treated with the oxidation reaction of Example 2, and subjected to the biotinylation conditions of Example 3.

Samples 1 and 2 had identical migration rates, demonstrating that the uncapped RNAs were not oxidized and biotinylated. Sample 3 migrated more slowly than Samples 1 and 2, while Sample 4 exhibited the slowest migration. The difference in migration of the RNAs in Samples 3 and 4 demonstrates that the capped RNAs were specifically biotinylated.

In some cases, mRNAs having intact 5′ ends may be enriched by binding the molecule containing a reactive amine group to a suitable solid phase substrate such as the inside of the vessel containing the mRNAs, magnetic beads, chromatography matrices, or nylon or nitrocellulose membranes. For example, where the molecule having a reactive amine group is biotin, the solid phase substrate may be coupled to avidin or streptavidin. Alternatively, where the molecule having the reactive amine group is an antibody or receptor ligand, the solid phase substrate may be coupled to the cognate antigen or receptor. Finally, where the molecule having a reactive amine group comprises an oligonucleotide, the solid phase substrate may comprise a complementary oligonucleotide.

The mRNAs having intact 5′ ends may be released from the solid phase following the enrichment procedure. For example, where the dialdehyde is coupled to biotin hydrazide and the solid phase comprises streptavidin, the mRNAs may be released from the solid phase by simply heating to 95 degrees Celsius in 2% SDS. In some methods, the molecule having a reactive amine group may also be cleaved from the mRNAs having intact 5′ ends following enrichment. Example 5 describes the capture of biotinylated mRNAs with streptavidin coated beads and the release of the biotinylated mRNAs from the beads following enrichment.

EXAMPLE 5

<u>Capture and Release of Biotinylated mRNAs Using Streptavidin Coated Beads</u>

The streptavidin coated magnetic beads were prepared according to the manufacturer's instructions (CPG Inc., USA). The biotinylated mRNAs were added to a

hybridization buffer (1.5 M NaCl, pH 5 - 6). After incubating for 30 minutes, the unbound and nonbiotinylated material was removed. The beads were then washed several times in water with 1% SDS. The beads thus obtained were incubated for 15 minutes at 95°C in water containing 2% SDS.

Example 6 demonstrates the efficiency with which biotinylated mRNAs were recovered from the streptavidin coated beads.

EXAMPLE 6

Efficiency of Recovery of Biotinylated mRNAs

The efficiency of the recovery procedure was evaluated as follows. Capped RNAs were labeled with ^{32}pCp, oxidized, biotinylated and bound to streptavidin coated beads as described above. Subsequently, the bound RNAs were incubated for 5, 15 or 30 minutes at 95°C in the presence of 2% SDS.

The products of the reaction were analyzed by electrophoresis on 12% polyacrylamide gels under denaturing conditions (7 M urea). The gels were subjected to autoradiography. During this manipulation, the hydrazone bonds were not reduced.

Increasing amounts of nucleic acids were recovered as incubation times in 2% SDS increased, demonstrating that biotinylated mRNAs were efficiently recovered.

In an alternative method for obtaining mRNAs having intact 5′ ends, an oligonucleotide which has been derivatized to contain a reactive amine group is specifically coupled to mRNAs having an intact cap. Preferably, the 3′ end of the mRNA is blocked prior to the step in which the aldehyde groups are joined to the derivatized oligonucleotide, as described above, so as to prevent the derivatized oligonucleotide from being joined to the 3′ end of the mRNA. For example, pCp may be attached to the 3′ end of the mRNA using T4 RNA ligase as described in example 1. However, as discussed above, blocking the 3′ end of the mRNA is an optional step. Derivatized oligonucleotides may be prepared as described in Example 7.

EXAMPLE 7

Derivatization of Oligonucleotides

An oligonucleotide phosphorylated at its 3′ end was converted to a 3′ hydrazide in 3′ by treatment with an aqueous solution of hydrazine or of dihydrazide of the formula $H_2N(R1)NH_2$ at about 1 to 3 M, and at pH 4.5 at a temperature of 8°C overnight. This incubation was performed in the presence of a carbodiimide type agent soluble in water such as 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide at a final concentration of 0.3 M.

The derivatized oligonucleotide was then separated from the other agents and products using a standard technique for isolating oligonucleotides.

As discussed above, the mRNAs to be enriched may be treated to eliminate the 3′ OH groups which may be present thereon. This may be accomplished by enzymatic ligation of sequences lacking a 3′ OH, such as pCp, as described in Example 1. Alternatively, the 3′ OH groups may be eliminated by alkaline hydrolysis as described in Example 8 below.

EXAMPLE 8

Elimination of 3' OH Groups of mRNA Using Alkaline Hydrolysis

In a total volume of 100 µl of 0.1 N sodium hydroxide, 1.5 µg mRNA is incubated for 40 to 60 minutes at 4°C. The solution is neutralized with acetic acid and precipitated with ethanol.

Following the optional elimination of the 3′ OH groups, the diol groups at the 5′ ends of the mRNAs are oxidized as described below in Example 9.

EXAMPLE 9

Oxidation of Diols of mRNA

Up to 1 OD unit of RNA was dissolved in 9 µl of buffer (0.1 M sodium acetate, pH 6-7) or water and 3 µl of freshly prepared 0.1 M sodium periodate solution. The reaction was incubated for 1 h in the dark at 4°C or room temperature. Following the incubation, the reaction was stopped by adding 4 µl of 10% ethylene glycol. Thereafter the mixture was incubated at room temperature for 15 minutes. After ethanol precipitation, the product was resuspended in at least 10 µl of water or appropriate buffer and dialyzed against water.

Following oxidation of the diol groups at the 5′ ends of the mRNAs, the derivatized oligonucleotide was joined to the resulting aldehydes as described in Example 10.

EXAMPLE 10

Ligature of Aldehydes of mRNA to Derivatized Oligonucleotides

The oxidized mRNA was dissolved in an acidic medium such as 50 µl of sodium acetate pH 4-6. Fifty µl of a solution of the derivatized oligonucleotide were added in order to obtain an mRNA:derivatized oligonucleotide ratio of 1:20. The mixture was reduced with a borohydride and incubated for 2 h at 37°C or overnight (14 h) at 10°C. The mixture was then ethanol precipitated, resuspended in 10 µl or more of water or appropriate buffer and dialyzed against distilled water. If desired, the resulting product may be analyzed using acrylamide gel electrophoresis, HPLC analysis, or other conventional techniques.

Following the attachment of the derivatized oligonucleotide to the mRNAs, a reverse transcription reaction may be performed as described in Example 11 below.

EXAMPLE 11

Reverse Transcription of mRNAs Ligatured to Derivatized Oligonucleotides

An oligodeoxyribonucleotide was derivatized as follows. Three OD units of an oligodeoxyribonucleotide of sequence 5'ATCAAGAATTCGCACGAGACCATTA3' (SEQ ID NO:3) having 5′-OH and 3′-P ends were dissolved in 70 µl of a 1.5 M hydroxybenzotriazole solution, pH 5.3, prepared in dimethylformamide/water (75:25) containing 2 µg of 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide. The mixture was incubated for 2 h 30 min at 22°C and then precipitated twice in $LiClO_4$/acetone. The pellet was resuspended in 200 µl of 0.25 M hydrazine and incubated at 8°C from 3 to 14 h. Following the hydrazine reaction, the mixture was precipitated twice in $LiClO_4$/acetone.

The messenger RNAs to be reverse transcribed were extracted from blocks of placenta having sides of 2 cm which had been stored at -80°C. The total RNA was extracted using conventional acidic phenol techniques. Oligo-dT chromatography was used to purify the mRNAs. The integrity of the mRNAs was checked by Northern-blotting.

The diol groups on 7 μg of the placental mRNAs were oxidized as described above in Example 9. The derivatized oligonucleotide was joined to the mRNAs as described in Example 10 above except that the precipitation step was replaced by an exclusion chromatography step to remove derivatized oligodeoxyribonucleotides which were not joined to mRNAs. Exclusion chromatography was performed as follows:

Ten ml of Ultrogel AcA34 (BioSepra#230151) gel, a mix of agarose and acrylamide, were equilibrated in 50 ml of a solution of 10 mM Tris pH 8.0, 300 mM NaCl, 1 mM EDTA, and 0.05% SDS. The mixture was allowed to sediment. The supernatant was eliminated and the gel was resuspended in 50 ml of buffer. This procedure was repeated 2 or 3 times.

A glass bead (diameter 3 mm) was introduced into a 2 ml disposable pipette (length 25 cm). The pipette was filled with the gel suspension until the height of the gel stabilized at 1 cm from the top of the pipette. The column was then equilibrated with 20 ml of equilibration buffer (10 mM Tris HCl pH 7.4, 20 mM NaCl).

Ten μl of the mRNA which had reacted with the derivatized oligonucleotide were mixed in 39 μl of 10 mM urea and 2 μl of blue-glycerol buffer, which had been prepared by dissolving 5 mg of bromophenol blue in 60% glycerol (v/v), and passing the mixture through a 0.45 μm diameter filter.

The column was then loaded with the mRNAs coupled to the oligonucleotide. As soon as the sample had penetrated, equilibration buffer was added. Hundred μl fractions were then collected. Derivatized oligonucleotide which had not been attached to mRNA appeared in fraction 16 and later fractions. Thus, fractions 3 to 15 were combined and precipitated with ethanol.

To determine whether the derivatized oligonucleotide was actually linked to mRNA, one tenth of the combined fractions were spotted twice on a nylon membrane and hybridized to a radioactive probe using conventional techniques. The ^{32}P labeled probe used in these hybridizations was an oligodeoxyribonucleotide of sequence 5'TAATGGTCTCGTGCGAATTCTTGAT3' (SEQ ID NO:4) anticomplementary to the derivatized oligonucleotide. A signal observed after autoradiography, indicated that the derivatized oligonucleotide had been truly joined to the mRNA.

The remaining nine tenth of the mRNAs which had reacted with the derivatized oligonucleotide was reverse transcribed as follows. A reverse transcription reaction was

carried out with reverse transcriptase following the manufacturer's instructions and 50 pmol of nonamers with random sequence as primers.

To ensure that reverse transcription had been carried out through the cap structure, two types of experiments were performed.

In the first approach, after elimination of RNA of the cDNA:RNA heteroduplexes obtained from the reverse transcription reaction by an alkaline hydrolysis, a portion of the resulting single stranded cDNAs was spotted on a positively charged membrane and hybridized, using conventional methods, to a ^{32}P labeled probe having a sequence identical to that of the derivatized oligonucleotide. Control spots containing, 1 pmol, 100 fmol, 50 fmol, 10 fmol and 1 fmol of a control oligodeoxyribonucleotide of sequence identical to that of the derivatized oligonucleotide were included. The signal observed in the spots containing the cDNA indicated that approximately 15 fmol of the derivatized oligonucleotide had been reverse transcribed. These results demonstrate that the reverse transcription can be performed through the cap and, in particular, that reverse transcriptase crosses the 5′-P-P-P-5′ bond of the cap of eukaryotic messenger RNAs.

In the second type of experiment, the single stranded cDNAs obtained from the above first strand synthesis were used as template for PCR reactions. Two types of reactions were carried out. First, specific amplification of the mRNAs for alpha globin, dehydrogenase, pp15 and elongation factor E4 were carried out using the following pairs of oligodeoxyribonucleotide primers.

alpha-globin

GLO-S: 5'CCG ACA AGA CCA ACG TCA AGG CCG C3' (SEQ ID NO:5)

GLO-As: 5'TCA CCA GCA GGC AGT GGC TTA GGA G 3′ (SEQ ID NO:6)

dehydrogenase

3 DH-S: 5'AGT GAT TCC TGC TAC TTT GGA TGG C3' (SEQ ID NO:7)

3 DH-As: 5'GCT TGG TCT TGT TCT GGA GTT TAG A3' (SEQ ID NO:8)

pp15

PP15-S: 5'TCC AGA ATG GGA GAC AAG CCA ATT T3' (SEQ ID NO:9)

PP15-As: 5'AGG GAG GAG GAA ACA GCG TGA GTC C3' (SEQ ID NO:10)

Elongation factor E4

EFA1-S: 5'ATG GGA AAG GAA AAG ACT CAT ATC A3' (SEQ ID NO:11)

EF1A-As: 5'AGC AGC AAC AAT CAG GAC AGC ACA G3' (SEQ ID NO:12)

Second, non specific amplifications were also carried out with the antisense oligodeoxyribonucleotides of the pairs described above and with a primer derived from the sequence of the derivatized oligodeoxyribonucleotide (5'ATCAAGAATTCGCACGAGACCATTA3') (SEQ ID NO:13).

One twentieth of the following RT-PCR product samples were run on a 1.5% agarose gel and stained with ethidium bromide.

Sample 1: The products of a PCR reaction using the globin primers of SEQ ID NOs 5 and 6 in the presence of cDNA.

Sample 2: The products of a PCR reaction using the globin primers of SEQ ID NOs 5 and 6 in the absence of added cDNA.

Sample 3: The products of a PCR reaction using the dehydrogenase primers of SEQ ID NOs 7 and 8 in the presence of cDNA.

Sample 4: The products of a PCR reaction using the dehydrogenase primers of SEQ ID NOs 7 and 8 in the absence of added cDNA.

Sample 5: The products of a PCR reaction using the pp15 primers of SEQ ID NOs 9 and 10 in the presence of cDNA.

Sample 6: The products of a PCR reaction using the pp15 primers of SEQ ID NOs 9 and 10 in the absence of added cDNA.

Sample 7: The products of a PCR reaction using the EIF4 primers of SEQ ID NOs 11 and 12 in the presence of added cDNA.

Sample 8: The products of a PCR reaction using the EIF4 primers of SEQ ID NOs 11 and 12 in the absence of added cDNA.

A band of the size expected for the PCR product was observed only in samples 1, 3, 5 and 7, thus indicating the presence of the corresponding sequence in the cDNA population.

PCR reactions were also carried out with the antisense oligonucleotides of the globin and dehydrogenase primers (SEQ ID NOs 6 and 8) and an oligonucleotide whose sequence corresponds to that of the derivatized oligonucleotide. The presence of PCR products of the expected size in the samples equivalent to above samples 1 and 3 indicated that the derivatized oligonucleotide had been linked to mRNA.

The above examples summarize the chemical procedure for enriching mRNAs for those having intact 5′ ends as illustrated in Figure 1. Further detail regarding the chemical approaches for obtaining such mRNAs are disclosed in International Application No. WO96/34981, published November 7, 1996, which is incorporated herein by reference. Strategies based on the above chemical modifications to the 5′ cap structure may be utilized to generate cDNAs selected to include the 5′ ends of the mRNAs from which they derived. In one version of such procedures, the 5′ ends of the mRNAs are modified as described above. Thereafter, a reverse transcription reaction is conducted to extend a primer complementary to the 5′ end of the mRNA. Single stranded RNAs are eliminated to obtain a population of cDNA/mRNA heteroduplexes in which the mRNA includes an intact 5′ end. The resulting heteroduplexes may be captured on a solid phase coated with a molecule capable of interacting with the molecule used to derivatize the 5′ end of the mRNA. Thereafter, the strands of the heteroduplexes are separated to recover single stranded first cDNA strands which include the 5′ end of the mRNA. Second strand cDNA synthesis may then proceed using conventional techniques. For example, the procedures disclosed in WO 96/34981 or in Carninci. *et al.*, *Genomics* 37:327-336, 1996, the disclosures of which are incorporated herein by reference, may be employed to select cDNAs which include the sequence derived from the 5′ end of the coding sequence of the mRNA.

Following ligation of the oligonucleotide tag to the 5′ cap of the mRNA, a reverse transcription reaction is conducted to extend a primer complementary to the mRNA to the 5′ end of the mRNA. Following elimination of the RNA component of the resulting heteroduplex using standard techniques, second strand cDNA synthesis is conducted with a primer complementary to the oligonucleotide tag.

2. Enzymatic Methods for Obtaining mRNAs having Intact 5′ Ends

Other techniques for selecting cDNAs extending to the 5′ end of the mRNA from which they are derived are fully enzymatic. Some versions of these techniques are disclosed in Dumas Milne Edwards J.B. (Doctoral Thesis of Paris VI University, Le clonage des ADNc complets: difficultes et perspectives nouvelles. Apports pour l'etude de la regulation de l'expression de la tryptophane hydroxylase de rat, 20 Dec. 1993), EP0 625572 and Kato *et al.*, *Gene* **150**:243-250, 1994, the disclosures of which are incorporated herein by reference.

Briefly, in such approaches, isolated mRNA is treated with alkaline phosphatase to remove the phosphate groups present on the 5′ ends of uncapped incomplete mRNAs. Following this procedure, the cap present on full length mRNAs is enzymatically removed with a decapping enzyme such as T4 polynucleotide kinase or tobacco acid pyrophosphatase. An oligonucleotide, which may be either a DNA oligonucleotide or a DNA-RNA hybrid oligonucleotide having RNA at its 3′ end, is then ligated to the phosphate present at the 5′ end of the decapped mRNA using T4 RNA ligase. The oligonucleotide may include a restriction site to facilitate cloning of the cDNAs following their synthesis. Example 12 below describes one enzymatic method based on the doctoral thesis of Dumas.

EXAMPLE 12

Enzymatic Approach for Obtaining 5′ ESTs

Twenty micrograms of PolyA+ RNA were dephosphorylated using Calf Intestinal Phosphatase (Biolabs). After a phenol chloroform extraction, the cap structure of mRNA was hydrolysed using the Tobacco Acid Pyrophosphatase (purified as described by Shinshi *et al.*., *Biochemistry* **15**: 2185-2190, 1976) and a hemi 5′DNA/RNA-3′ oligonucleotide having an unphosphorylated 5′ end, a stretch of adenosine ribophosphate at the 3′ end, and an EcoRI site near the 5′ end was ligated to the 5′P ends of mRNA using the T4 RNA ligase (Biolabs). Oligonucleotides suitable for use in this procedure are preferably 30 to 50 bases in length. Oligonucleotides having an unphosphorylated 5′ end may be synthesized by adding a fluorochrome at the 5′ end. The inclusion of a stretch of adenosine ribophosphates at the 3′ end of the oligonucleotide increases ligation efficiency. It will be appreciated that the oligonucleotide may contain cloning sites other than EcoRI.

Following ligation of the oligonucleotide to the phosphate present at the 5′ end of the decapped mRNA, first and second strand cDNA synthesis is carried out using conventional methods or those specified in EP0 625,572 and Kato *et al. supra,* and Dumas Milne Edwards, *supra,* the disclosures of which are incorporated herein by reference. The resulting cDNA may then be ligated into vectors such as those disclosed in Kato *et al., supra* or other nucleic acid vectors known to those skilled in the art using techniques such as those described in Sambrook *et al.*, Molecular Cloning: A Laboratory Manual 2d Ed., Cold Spring Harbor Laboratory Press, 1989, the disclosure of which is incorporated herein by reference.

II. Obtention and Characterization of the 5′ ESTs of the Present Invention

The 5' ESTs of the present invention were obtained using the aforementioned chemical and enzymatic approaches for enriching mRNAs for those having intact 5′ ends as decribed below.

1. Obtention of 5' ESTS Using mRNAs with Intact 5' Ends

First, mRNAs were prepared as described in Example 13 below.

EXAMPLE 13

Preparation of mRNA With Intact 5' Ends

Total human RNAs or polyA$^+$ RNAs derived from 29 different tissues were respectively purchased from LABIMO and CLONTECH and used to generate 44 cDNA libraries as follows. The purchased RNA had been isolated from cells or tissues using acid guanidium thiocyanate-phenol-chloroform extraction (Chomczyniski and Sacchi, *Analytical Biochemistry* **162**:156-159, 1987). PolyA$^+$ RNA was isolated from total RNA (LABIMO) by two passes of oligo dT chromatography, as described by Aviv and Leder, *Proc. Natl. Acad. Sci. USA* **69**:1408-1412, 1972 in order to eliminate ribosomal RNA.

The quality and the integrity of the polyA+ RNAs were checked. Northern blots hybridized with a globin probe were used to confirm that the mRNAs were not degraded. Contamination of the polyA$^+$ mRNAs by ribosomal sequences was checked using Northern blots and a probe derived from the sequence of the 28S rRNA. Preparations of mRNAs with

less than 5% of rRNAs were used in library construction. To avoid constructing libraries with RNAs contaminated by exogenous sequences (prokaryotic or fungal), the presence of bacterial 16S ribosomal sequences or of two highly expressed fungal mRNAs was examined using PCR.

Following preparation of the mRNAs, the above described chemical and/or the enzymatic procedures for enriching mRNAs for thoses having intact 5′ ends were employed to obtain 5′ ESTs from various tissues. In both approaches, an oligonucleotide tag was attached to the 5′ ends of the mRNAs. The oligonucleotide tag had an EcoRI site therein to facilitate later cloning procedures. To facilitate the processing of single stranded and double stranded cDNA obtained in the construction of the librairies, the same nucleotidic sequence was used to design the ligated oligonucleotide in both chemical and enzymatic approaches. Nevertheless, in the chemical procedure, the tag used was an oligodeoxyribonucleotide which was linked to the cap of the mRNA whereas in the enzymatic ligation, the tag was a chimeric hemi 5'DNA/RNA3' oligonucleotide which was ligated to the 5' end of decapped mRNA as described in example 12.

Following attachment of the oligonucleotide tag to the mRNA by either the chemical or enzymatic methods, the integrity of the mRNA was examined by performing a Northern blot with 200 to 500 ng of mRNA using a probe complementary to the oligonucleotide tag before performing the first strand synthesis as described in example 14.

EXAMPLE 14

cDNA Synthesis Using mRNA Templates Having Intact 5′ Ends

For the mRNAs joined to oligonucleotide tags using both the chemical and enzymatic methods, first strand cDNA synthesis was performed using the Superscript II (Gibco BRL) or the Rnase H Minus M-MLV (Promega) reverse transcriptase with random nonamers as primers. In order to protect internal EcoRI sites in the cDNA from digestion at later steps in the procedure, methylated dCTP was used for first strand synthesis. After removal of RNA by an alkaline hydrolysis, the first strand of cDNA was precipitated using isopropanol in order to eliminate residual primers.

For both the chemical and the enzymatic methods, the second strand of the cDNA was synthesized with a Klenow fragment using a primer corresponding to the 5′ end of the

ligated oligonucleotide described in Example 12. Preferably, the primer is 20-25 bases in length. Methylated dCTP was also used for second strand synthesis in order to protect internal EcoRI sites in the cDNA from digestion during the cloning process.

Following cDNA synthesis, the cDNAs were cloned into pBlueScript as described in Example 15 below.

EXAMPLE 15

Cloning of cDNAsderived from mRNA with intact 5' ends into BlueScript

Following second strand synthesis, the ends of the cDNA were blunted with T4 DNA polymerase (Biolabs) and the cDNA was digested with EcoRI. Since methylated dCTP was used during cDNA synthesis, the EcoRI site present in the tag was the only hemi-methylated site, hence the only site susceptible to EcoRI digestion. The cDNA was then size fractionated using exclusion chromatography (AcA, Biosepra) and fractions corresponding to cDNAs of more than 150 bp were pooled and ethanol precipitated. The cDNA was directionally cloned into the SmaI and EcoRI ends of the phagemid pBlueScript vector (Stratagene). The ligation mixture was electroporated into bacteria and propagated under appropriate antibiotic selection.

Clones containing the oligonucleotide tag attached were then selected as described in Example 16 below.

EXAMPLE 16

Selection of Clones Having the Oligonucleotide Tag Attached Thereto

The plasmid DNAs containing 5′ EST libraries made as described above were purified (Qiagen). A positive selection of the tagged clones was performed as follows. Briefly, in this selection procedure, the plasmid DNA was converted to single stranded DNA using gene II endonuclease of the phage F1 in combination with an exonuclease (Chang *et al.*, *Gene* **127**:95-8, 1993) such as exonuclease III or T7 gene 6 exonuclease. The resulting single stranded DNA was then purified using paramagnetic beads as described by Fry *et al.*, *Biotechniques*, **13**: 124-131, 1992. In this procedure, the single stranded DNA was hybridized with a biotinylated oligonucleotide having a sequence corresponding to the 3′ end of the oligonucleotide described in Example 13. Preferably, the primer has a length of 20-25

bases. Clones including a sequence complementary to the biotinylated oligonucleotide were captured by incubation with streptavidin coated magnetic beads followed by magnetic selection. After capture of the positive clones, the plasmid DNA was released from the magnetic beads and converted into double stranded DNA using a DNA polymerase such as the ThermoSequenase obtained from Amersham Pharmacia Biotech. Alternatively, protocoles such as the one described in the Gene Trapper kit available from Gibco BRL may be used. The double stranded DNA was then electroporated into bacteria. The percentage of positive clones having the 5′ tag oligonucleotide was estimated to typically rank between 90 and 98% using dot blot analysis.

Following electroporation, the libraries were ordered in 384-microtiter plates (MTP). A copy of the MTP was stored for future needs. Then the libraries were transferred into 96 MTP and sequenced as described below.

EXAMPLE 17

Sequencing of Inserts in Selected Clones

Plasmid inserts were first amplified by PCR on PE 9600 thermocyclers (Perkin-Elmer, Applied Biosystems Division, Foster City, CA), using standard SETA-A and SETA-B primers (Genset SA), AmpliTaqGold (Perkin-Elmer), dNTPs (Boehringer), buffer and cycling conditions as recommended by the Perkin-Elmer Corporation.

PCR products were then sequenced using automatic ABI Prism 377 sequencers (Perkin Elmer). Sequencing reactions were performed using PE 9600 thermocyclers with standard dye-primer chemistry and ThermoSequenase (Amersham Pharmacia Biotech). The primers used were either T7 or 21M13 (available from Genset SA) as appropriate. The primers were labeled with the JOE, FAM, ROX and TAMRA dyes. The dNTPs and ddNTPs used in the sequencing reactions were purchased from Boehringer. Sequencing buffer, reagent concentrations and cycling conditions were as recommended by Amersham.

Following the sequencing reaction, the samples were precipitated with ethanol, resuspended in formamide loading buffer, and loaded on a standard 4% acrylamide gel. Electrophoresis was performed for 2.5 hours at 3000V on an ABI 377 sequencer, and the sequence data were collected and analyzed using the ABI Prism DNA Sequencing Analysis Software, version 2.1.2.

2. Computer analysis of the Obtained 5' ESTs: Construction of NetGene and SignalTag databases

The sequence data from the 44 cDNA libraries made as described above were transferred to a proprietary database, where quality control and validation steps were performed. A proprietary base-caller, working using a Unix system, automatically flagged suspect peaks, taking into account the shape of the peaks, the inter-peak resolution, and the noise level. The proprietary base-caller also performed an automatic trimming. Any stretch of 25 or fewer bases having more than 4 suspect peaks was considered unreliable and was discarded. Sequences corresponding to cloning vector or ligation oligonucleotides were automatically removed from the EST sequences. However, the resulting EST sequences may contain 1 to 5 bases belonging to the above mentioned sequences at their 5′ end. If needed, these can easily be removed on a case to case basis.

Following sequencing as described above, the sequences of the 5′ ESTs were entered in NetGene™, a proprietary database called for storage and manipulation as described below. It will be appreciated by those skilled in the art that the data could be stored and manipulated on any medium which can be read and accessed by a computer. Computer readable media include magnetically, optically, or electronically readable media. For example, the computer readable media may be a hard disc, a floppy disc, a magnetic tape, CD-ROM, RAM, or ROM as well as other types of other media known to those skilled in the art.

In addition, the sequence data may be stored and manipulated in a variety of data processor programs in a diversity of formats. For instance, the sequence data may be stored as text in a word processing file, such as Microsoft WORD or WORDPERFECT or as an ASCII file in a variety of database programs familiar to those of skill in the art, such as DB2, SYBASE, or ORACLE.

The computer readable media on which the sequence information is stored may be in a personal computer, a network, a server or other computer systems known to those skilled in the art. The computer or other system preferably includes the storage media described above, and a processor for accessing and manipulating the sequence data. Once the sequence data has been stored, it may be manipulated and searched to locate those stored sequences which contain a desired nucleic acid sequence or which encode a protein having a particular functional domain. For example, the stored sequence information may be compared to other

known sequences to identify homologies, motifs implicated in biological function, or structural motifs.

Programs which may be used to search or compare the stored sequences include the MacPattern (EMBL), BLAST, and BLAST2 program series (NCBI), basic local alignment search tool programs for nucleotide (BLASTN) and peptide (BLASTX) comparisons (Altschul *et al*, *J. Mol. Biol.* **215**: 403, 1990) and FASTA (Pearson and Lipman, *Proc. Natl. Acad. Sci. USA* **85**: 2444, 1988). The BLAST programs then extend the alignments on the basis of defined match and mismatch criteria.

Motifs which may be detected using the above programs and those described in Example 28 include sequences encoding leucine zippers, helix-turn-helix motifs, glycosylation sites, ubiquitination sites, alpha helices, and beta sheets, signal sequences encoding signal peptides which direct the secretion of the encoded proteins, sequences implicated in transcription regulation such as homeoboxes, acidic stretches, enzymatic active sites, substrate binding sites, and enzymatic cleavage sites.

Before searching the cDNAs in the NetGene™ database for sequence motifs of interest, cDNAs derived from mRNAs which were not of interest were identified and eliminated from further consideration as described in Example 18 below.

EXAMPLE 18

Elimination of Undesired Sequences from Further Consideration

5′ ESTs in the NetGene™ database which were derived from undesired sequences such as transfer RNAs, ribosomal RNAs, mitochondrial RNAs, prokaryotic RNAs, fungal RNAs, Alu sequences, L1 sequences, or repeat sequences were identified using the FASTA and BLASTN programs with the parameters listed in Table I.

To eliminate 5′ ESTs encoding tRNAs from further consideration, the 5′ EST sequences were compared to the sequences of 1190 known tRNAs obtained from EMBL release 38, of which 100 were human. The comparison was performed using FASTA on both strands of the 5′ ESTs. Sequences having more than 80% homology over more than 60 nucleotides were identified as tRNA. Of the 144,341 sequences screened, 26 were identified as tRNAs and eliminated from further consideration.

To eliminate 5′ ESTs encoding rRNAs from further consideration, the 5′ EST sequences were compared to the sequences of 2497 known rRNAs obtained from EMBL release 38, of which 73 were human. The comparison was performed using BLASTN on both strands of the 5′ ESTs with the parameter S=108. Sequences having more than 80% homology over stretches longer than 40 nucleotides were identified as rRNAs. Of the 144,341 sequences screened, 3,312 were identified as rRNAs and eliminated from further consideration.

To eliminate 5′ ESTs encoding mtRNAs from further consideration, the 5′ EST sequences were compared to the sequences of the two known mitochondrial genomes for which the entire genomic sequences are available and all sequences transcribed from these mitochondrial genomes including tRNAs, rRNAs, and mRNAs for a total of 38 sequences. The comparison was performed using BLASTN on both strands of the 5′ ESTs with the parameter S=108. Sequences having more than 80% homology over stretches longer than 40 nucleotides were identified as mtRNAs. Of the 144,341 sequences screened, 6,110 were identified as mtRNAs and eliminated from further consideration.

Sequences which might have resulted from exogenous contaminants were eliminated from further consideration by comparing the 5′ EST sequences to release 46 of the EMBL bacterial and fungal divisions using BLASTN with the parameter S=144. All sequences having more than 90% homology over at least 40 nucleotides were identified as exogenous contaminants. Of the 42 cDNA libraries examined, the average percentages of prokaryotic and fungal sequences contained therein were 0.2% and 0.5% respectively. Among these sequences, only one could be identified as a sequence specific to fungi. The others were either fungal or prokaryotic sequences having homologies with vertebrate sequences or including repeat sequences which had not been masked during the electronic comparison.

In addition, the 5′ ESTs were compared to 6093 Alu sequences and 1115 L1 sequences to mask 5′ ESTs containing such repeat sequences. 5′ ESTs including THE and MER repeats, SSTR sequences or satellite, micro-satellite, or telomeric repeats were also eliminated from further consideration. On average, 11.5% of the sequences in the libraries contained repeat sequences. Of this 11.5%, 7% contained Alu repeats, 3.3% contained L1 repeats and the remaining 1.2% were derived from the other screened types of repetitive sequences. These percentages are consistent with those found in cDNA libraries prepared by

other groups. For example, the cDNA libraries of Adams *et al.* contained between 0% and 7.4% Alu repeats depending on the source of the RNA which was used to prepare the cDNA library (Adams *et al.*, *Nature* **377**:174, 1996).

The sequences of those 5′ ESTs remaining after the elimination of undesirable sequences were compared with the sequences of known human mRNAs to determine the accuracy of the sequencing procedures described above.

EXAMPLE 19

Measurement of Sequencing Accuracy by Comparison to Known Sequences

To further determine the accuracy of the sequencing procedure described above, the sequences of 5′ ESTs derived from known sequences were identified and compared to the original known sequences. First, a FASTA analysis with overhangs shorter than 5 bp on both ends was conducted on the 5′ ESTs to identify those matching an entry in the public human mRNA database. The 6655 5′ ESTs which matched a known human mRNA were then realigned with their cognate mRNA and dynamic programming was used to include substitutions, insertions, and deletions in the list of "errors" which would be recognized. Errors occurring in the last 10 bases of the 5′ EST sequences were ignored to avoid the inclusion of spurious cloning sites in the analysis of sequencing accuracy.

This analysis revealed that the sequences incorporated in the NetGene™ database had an accuracy of more than 99.5%.

To determine the efficiency with which the above selection procedures select cDNAs which include the 5′ ends of their corresponding mRNAs, the following analysis was performed.

EXAMPLE 20

Determination of Efficiency of 5′ EST Selection

To determine the efficiency at which the above selection procedures isolated 5′ ESTs which included sequences close to the 5′ end of the mRNAs from which they derived, the sequences of the ends of the 5′ ESTs derived from the elongation factor 1 subunit α and

ferritin heavy chain genes were compared to the known cDNA sequences of these genes. Since the transcription start sites of both genes are well characterized, they may be used to determine the percentage of derived 5′ ESTs which included the authentic transcription start sites.

For both genes, more than 95% of the obtained 5' ESTs actually included sequences close to or upstream of the 5′ end of the corresponding mRNAs.

To extend the analysis of the reliability of the procedures for isolating 5′ ESTs from ESTs in the NetGene™ database, a similar analysis was conducted using a database composed of human mRNA sequences extracted from GenBank database release 97 for comparison. The 5' ends of more than 85% of 5′ ESTs derived from mRNAs included in the GeneBank database were located close to the 5′ ends of the known sequence. As some of the mRNA sequences available in the GenBank database are deduced from genomic sequences, a 5′ end matching with these sequences will be counted as an internal match. Thus, the method used here underestimates the yield of ESTs including the authentic 5′ ends of their corresponding mRNAs.

The EST libraries made above included multiple 5′ ESTs derived from the same mRNA. The sequences of such 5′ ESTs were compared to one another and the longest 5′ ESTs for each mRNA were identified. Overlapping cDNAs were assembled into continuous sequences (contigs). The resulting continuous sequences were then compared to public databases to gauge their similarity to known sequences, as described in Example 21 below.

EXAMPLE 21

Clustering of the 5′ ESTs and Calculation of Novelty Indices for cDNA Libraries

For each sequenced EST library, the sequences were clustered by the 5′ end. Each sequence in the library was compared to the others with BLASTN2 (direct strand, parameters S=107). ESTs with High Scoring Segment Pairs (HSPs) at least 25 bp long, having 95% identical bases and beginning closer than 10 bp from each EST 5′ end were grouped. The longest sequence found in the cluster was used as representative of the group. A global clustering between libraries was then performed leading to the definition of super-contigs.

To assess the yield of new sequences within the EST libraries, a novelty rate (NR) was defined as: NR= 100 X (Number of new unique sequences found in the library/Total number of sequences from the library). Typically, novelty rating ranged between 10% and 41% depending on the tissue from which the EST library was obtained. For most of the libraries, the random sequencing of 5′ EST libraries was pursued until the novelty rate reached 20%.

Following characterization as described above, the collection of 5′ ESTs in NetGene™ was screened to identify those 5′ ESTs bearing potential signal sequences as described in Example 22 below.

EXAMPLE 22

Identification of Potential Signal Sequences in 5′ ESTs

The 5′ ESTs in the NetGene™ database were screened to identify those having an uninterrupted open reading frame (ORF) longer than 45 nucleotides beginning with an ATG codon and extending to the end of the EST. Approximately half of the cDNA sequences in NetGene™ contained such an ORF. The ORFs of these 5′ ESTs were then searched to identify potential signal motifs using slight modifications of the procedures disclosed in Von Heijne, *Nucleic Acids Res.* **14**:4683-4690, 1986, the disclosure of which is incorporated herein by reference. Those 5′ EST sequences encoding a stretch of at least 15 amino acid long with a score of at least 3.5 in the Von Heijne signal peptide identification matrix were considered to possess a signal sequence. Those 5′ ESTs which matched a known human mRNA or EST sequence and had a 5′ end more than 20 nucleotides downstream of the known 5′ end were excluded from further analysis. The remaining cDNAs having signal sequences therein were included in a database called SignalTag™.

To confirm the accuracy of the above method for identifying signal sequences, the analysis of Example 23 was performed.

EXAMPLE 23

Confirmation of Accuracy of Identification of Potential Signal Sequences in 5′ ESTs

The accuracy of the above procedure for identifying signal sequences encoding signal peptides was evaluated by applying the method to the 43 amino acids located at the N terminus of all human SwissProt proteins. The computed Von Heijne score for each protein was compared with the known characterization of the protein as being a secreted protein or a non-secreted protein. In this manner, the number of non-secreted proteins having a score higher than 3.5 (false positives) and the number of secreted proteins having a score lower than 3.5 (false negatives) could be calculated.

Using the results of the above analysis, the probability that a peptide encoded by the 5′ region of the mRNA is in fact a genuine signal peptide based on its Von Heijne's score was calculated based on either the assumption that 10% of human proteins are secreted or the assumption that 20% of human proteins are secreted. The results of this analysis are shown in Figure 2 and table IV.

Using the above method of identification of secretory proteins, 5′ ESTs of the following polypeptides known to be secreted were obtained: human glucagon, gamma interferon induced monokine precursor, secreted cyclophilin-like protein, human pleiotropin, and human biotinidase precursor. Thus, the above method successfully identified those 5′ ESTs which encode a signal peptide.

To confirm that the signal peptide encoded by the 5′ ESTs actually functions as a signal peptide, the signal sequences from the 5′ ESTs may be cloned into a vector designed for the identification of signal peptides. Such vectors are designed to confer the ability to grow in selective medium only to host cells containing a vector with an operably linked signal sequence. For example, to confirm that a 5′ EST encodes a genuine signal peptide, the signal sequence of the 5′ EST may be inserted upstream and in frame with a non-secreted form of the yeast invertase gene in signal peptide selection vectors such as those described in U.S. Patent No. 5,536,637, the disclosure of which is incorporated herein by reference. Growth of host cells containing signal sequence selection vectors with the correctly inserted 5′ EST signal sequence confirms that the 5′ EST encodes a genuine signal peptide.

Alternatively, the presence of a signal peptide may be confirmed by cloning the extended cDNAs obtained using the ESTs into expression vectors such as pXT1 (as described below in example 30), or by constructing promoter-signal sequence-reporter gene vectors which encode fusion proteins between the signal peptide and an assayable reporter protein. After introduction of these vectors into a suitable host cell, such as COS cells or NIH 3T3 cells, the growth medium may be harvested and analyzed for the presence of the secreted protein. The medium from these cells is compared to the medium from control cells containing vectors lacking the signal sequence or extended cDNA insert to identify vectors which encode a functional signal peptide or an authentic secreted protein.

Those 5′ ESTs which encoded a signal peptide, as determined by the method of Example 22 above, were further grouped into four categories based on their homology to known sequences as described in Example 24 below.

EXAMPLE 24

Categorization of 5′ ESTs Encoding a Signal Peptide

Those 5′ ESTs having a sequence not matching any known vertebrate sequence nor any publicly available EST sequence were designated "new." Of the sequences in the SignalTag™ database, 947 of the 5′ ESTs having a Von Heijne's score of at least 3.5 fell into this category.

Those 5′ ESTs having a sequence not matching any vertebrate sequence but matching a publicly known EST were designated "EST-ext", provided that the known EST sequence was extended by at least 40 nucleotides in the 5′ direction. Of the sequences in the SignalTag™ database, 150 of the 5′ ESTs having a Von Heijne's score of at least 3.5 fell into this category.

Those ESTs not matching any vertebrate sequence but matching a publicly known EST without extending the known EST by at least 40 nucleotides in the 5′ direction were designated "EST." Of the sequences in the SignalTag™ database, 599 of the 5′ ESTs having a Von Heijne's score of at least 3.5 fell into this category.

Those 5′ ESTs matching a human mRNA sequence but extending the known sequence by at least 40 nucleotides in the 5′ direction were designated "VERT-ext." Of the sequences in the SignalTag™ database, 23 of the 5′ ESTs having a Von Heijne's score of at

least 3.5 fell into this category. Included in this category was a 5′ EST which extended the known sequence of the human translocase mRNA by more than 200 bases in the 5′ direction. A 5′ EST which extended the sequence of a human tumor suppressor gene in the 5′ direction was also identified.

Table V shows the distribution of 5′ ESTs in each category and the number of 5′ ESTs in each category having a given minimum von Heijne's score.

3. Evaluation of Spatial and Temporal Expression of mRNAs Corresponding to the 5'ESTs or Extended cDNAs

Each of the 5′ ESTs was also categorized based on the tissue from which its corresponding mRNA was obtained, as described below in Example 25.

EXAMPLE 25

Categorization of Expression Patterns

Table VI shows the distribution of 5' ESTs in each of the above defined category with respect to the tissue from which the 5'ESTs of the corresponding mRNA were obtained.

Table II provides the sequence identification numbers of 5′ EST sequences derived from brain, the categories in which these sequences fall, and the von Heijne's score of the signal peptides which they encode. The 5′ EST sequences and the amino acid sequences they encode are provided in the appended sequence listings. Table III provides the sequence ID numbers of the 5′ ESTs and the sequences of the signal peptides which they encode. The sequences of the 5′ ESTs and the polypeptides they encode are provided in the sequence listing appended hereto.

The sequences of DNA SEQ ID NOs: 38-195 can readily be screened for any errors therein and any sequence ambiguities can be resolved by resequencing a fragment containing such errors or amibiguities on both strands. Such fragments may be obtained from the plasmids stored in the inventors' laboratory or can be isolated using the techniques described herein. Resolution of any such ambiguities or errors may be facilitated by using primers which hybridize to sequences located close to the ambiguous or erroneous sequences. For example, the primers may hybridize to sequences within 50-75 bases of the amibiguity or

error. Upon resolution of an error or ambiguity, the corresponding corrections can be made in the protein sequences encoded by the DNA containing the error or amibiguity.

In addition to categorizing the 5′ ESTs with respect to their tissue of origin, the spatial and temporal expression patterns of the mRNAs corresponding to the 5′ ESTs, as well as their expression levels, may be determined as described in Example 26 below. Characterization of the spatial and temporal expression patterns and expression levels of these mRNAs is useful for constructing expression vectors capable of producing a desired level of gene product in a desired spatial or temporal manner, as will be discussed in more detail below.

Furthermore, 5′ ESTs whose corresponding mRNAs are associated with disease states may also be identified. For example, a particular disease may result from the lack of expression, over expression, or under expression of an mRNA corresponding to a 5′ EST. By comparing mRNA expression patterns and quantities in samples taken from healthy individuals with those from individuals suffering from a particular disease, 5′ ESTs responsible for the disease may be identified.

It will be appreciated that the results of the above characterization procedures for 5′ ESTs also apply to extended cDNAs (obtainable as described below) which contain sequences adjacent to the 5′ ESTs. It will also be appreciated that if desired, characterization may be delayed until extended cDNAs have been obtained rather than characterizing the ESTs themselves.

EXAMPLE 26

Evaluation of Expression Levels and Patterns of mRNAs Corresponding to 5′ ESTs or Extended cDNAs

Expression levels and patterns of mRNAs corresponding to 5′ ESTs or extended cDNAs (obtainable as described below in example 27) may be analyzed by solution hybridization with long probes as described in International Patent Application No. WO 97/05277, the entire contents of which are hereby incorporated by reference. Briefly, a 5′ EST, extended cDNA, or fragment thereof corresponding to the gene encoding the mRNA to be characterized is inserted at a cloning site immediately downstream of a bacteriophage (T3,

T7 or SP6) RNA polymerase promoter to produce antisense RNA. Preferably, the 5′ EST or extended cDNA has 100 or more nucleotides. The plasmid is linearized and transcribed in the presence of ribonucleotides comprising modified ribonucleotides (*i.e.* biotin-UTP and DIG-UTP). An excess of this doubly labeled RNA is hybridized in solution with mRNA isolated from cells or tissues of interest. The hybridizations are performed under standard stringent conditions (40-50°C for 16 hours in an 80% formamide, 0.4 M NaCl buffer, pH 7-8). The unhybridized probe is removed by digestion with ribonucleases specific for single-stranded RNA (*i.e.* RNases CL3, T1, Phy M, U2 or A). The presence of the biotin-UTP modification enables capture of the hybrid on a microtitration plate coated with streptavidin. The presence of the DIG modification enables the hybrid to be detected and quantified by ELISA using an anti-DIG antibody coupled to alkaline phosphatase.

The 5′ ESTs, extended cDNAs, or fragments thereof may also be tagged with nucleotide sequences for the serial analysis of gene expression (SAGE) as disclosed in UK Patent Application No. 2 305 241 A, the entire contents of which are incorporated by reference. In this method, cDNAs are prepared from a cell, tissue, organism or other source of nucleic acid for which gene expression patterns must be determined. The resulting cDNAs are separated into two pools. The cDNAs in each pool are cleaved with a first restriction endonuclease, called an anchoring enzyme, having a recognition site which is likely to be present at least once in most cDNAs. The fragments which contain the 5′ or 3′ most region of the cleaved cDNA are isolated by binding to a capture medium such as streptavidin coated beads. A first oligonucleotide linker having a first sequence for hybridization of an amplification primer and an internal restriction site for a so-called tagging endonuclease is ligated to the digested cDNAs in the first pool. Digestion with the second endonuclease produces short tag fragments from the cDNAs.

A second oligonucleotide having a second sequence for hybridization of an amplification primer and an internal restriction site is ligated to the digested cDNAs in the second pool. The cDNA fragments in the second pool are also digested with the tagging endonuclease to generate short tag fragments derived from the cDNAs in the second pool. The tags resulting from digestion of the first and second pools with the anchoring enzyme and the tagging endonuclease are ligated to one another to produce so-called ditags. In some embodiments, the ditags are concatamerized to produce ligation products containing from 2

to 200 ditags. The tag sequences are then determined and compared to the sequences of the 5′ ESTs or extended cDNAs to determine which 5′ ESTs or extended cDNAs are expressed in the cell, tissue, organism, or other source of nucleic acids from which the tags were derived. In this way, the expression pattern of the 5′ ESTs or extended cDNAs in the cell, tissue, organism, or other source of nucleic acids is obtained.

Quantitative analysis of gene expression may also be performed using arrays. As used herein, the term array means a one dimensional, two dimensional, or multidimensional arrangement of full length cDNAs (*i.e.* extended cDNAs which include the coding sequence for the signal peptide, the coding sequence for the mature protein, and a stop codon), extended cDNAs, 5′ ESTs or fragments thereof of sufficient length to permit specific detection of gene expression. Preferably, the fragments are at least 15 nucleotides in length. More preferably, the fragments are at least 100 nucleotide long. More preferably, the fragments are more than 100 nucleotides in length. In some embodiments, the fragments may be more than 500 nucleotide long.

For example, quantitative analysis of gene expression may be performed with full length cDNAs as defined below, extended cDNAs, 5′ ESTs, or fragments thereof in a complementary DNA microarray as described by Schena *et al.* (*Science* **270**:467-470, 1995; *Proc. Natl. Acad. Sci. U.S.A.* **93**:10614-10619, 1996). Full length cDNAs, extended cDNAs, 5′ ESTs or fragments thereof are amplified by PCR and arrayed from 96-well microtiter plates onto silylated microscope slides using high-speed robotics. Printed arrays are incubated in a humid chamber to allow rehydration of the array elements and rinsed, once in 0.2% SDS for 1 min, twice in water for 1 min and once for 5 min in sodium borohydride solution. The arrays are submerged in water for 2 min at 95°C, transferred into 0.2% SDS for 1 min, rinsed twice with water, air dried and stored in the dark at 25°C.

Cell or tissue mRNA is isolated or commercially obtained and probes are prepared by a single round of reverse transcription. Probes are hybridized to 1 cm^2 microarrays under a 14 x 14 mm glass coverslip for 6-12 hours at 60°C. Arrays are washed for 5 min at 25°C in low stringency wash buffer (1 x SSC/0.2% SDS), then for 10 min at room temperature in high stringency wash buffer (0.1 x SSC/0.2% SDS). Arrays are scanned in 0.1 x SSC using a fluorescence laser scanning device fitted with a custom filter set. Accurate differential

expression measurements are obtained by taking the average of the ratios of two independent hybridizations.

Quantitative analysis of the expression of genes may also be performed with full length cDNAs, extended cDNAs, 5′ ESTs, or fragments thereof in complementary DNA arrays as described by Pietu *et al.*. (*Genome Research* 6:492-503, 1996). The full length cDNAs, extended cDNAs, 5′ ESTs or fragments thereof are PCR amplified and spotted on membranes. Then, mRNAs originating from various tissues or cells are labeled with radioactive nucleotides. After hybridization and washing in controlled conditions, the hybridized mRNAs are detected by phospho-imaging or autoradiography. Duplicate experiments are performed and a quantitative analysis of differentially expressed mRNAs is then performed.

Alternatively, expression analysis of the 5′ ESTs or extended cDNAs can be done through high density nucleotide arrays as described by Lockhart *et al.* (*Nature Biotechnology* 14: 1675-1680, 1996) and Sosnowsky *et al.* (*Proc. Natl. Acad. Sci.* 94:1119-1123, 1997). Oligonucleotides of 15-50 nucleotides corresponding to sequences of the 5′ ESTs or extended cDNAs are synthesized directly on the chip (Lockhart *et al.*, *supra*) or synthesized and then addressed to the chip (Sosnowsky *et al.*, *supra*). Preferably, the oligonucleotides are about 20 nucleotides in length.

cDNA probes labeled with an appropriate compound, such as biotin, digoxigenin or fluorescent dye, are synthesized from the appropriate mRNA population and then randomly fragmented to an average size of 50 to 100 nucleotides. The said probes are then hybridized to the chip. After washing as described in Lockhart *et al*, *supra* and application of different electric fields (Sonowsky et *al*, *supra*.), the dyes or labeling compounds are detected and quantified. Duplicate hybridizations are performed. Comparative analysis of the intensity of the signal originating from cDNA probes on the same target oligonucleotide in different cDNA samples indicates a differential expression of the mRNA corresponding to the 5′ EST or extended cDNA from which the oligonucleotide sequence has been designed.

III. Use of 5′ ESTs to Clone Extended cDNAs and to Clone the Corresponding Genomic DNAs

Once 5′ ESTs which include the 5′ end of the corresponding mRNAs have been selected using the procedures described above, they can be utilized to isolate extended cDNAs which contain sequences adjacent to the 5′ ESTs. The extended cDNAs may include the entire coding sequence of the protein encoded by the corresponding mRNA, including the authentic translation start site, the signal sequence, and the sequence encoding the mature protein remaining after cleavage of the signal peptide. Such extended cDNAs are referred to herein as "full length cDNAs." Alternatively, the extended cDNAs may include only the sequence encoding the mature protein remaining after cleavage of the signal peptide, or only the sequence encoding the signal peptide.

Example 27 below describes a general method for obtaining extended cDNAs using 5' ESTs. Example 28 below provides experimental results, using the method explained in example 27, describing several extended cDNAs including the entire coding sequence and authentic 5′ end of the corresponding mRNA for several secreted proteins.

The methods of Examples 27, 28, and 29 can also be used to obtain extended cDNAs which encode less than the entire coding sequence of the secreted proteins encoded by the genes corresponding to the 5′ ESTs. In some embodiments, the extended cDNAs isolated using these methods encode at least 10 amino acids of one of the proteins encoded by the sequences of SEQ ID NOs: 38-195. In further embodiments, the extended cDNAs encode at least 20 amino acids of the proteins encoded by the sequences of SEQ ID NOs: 38-195. In further embodiments, the extended cDNAs encode at least 30 amino amino acids of the sequences of SEQ ID NOs: 38-195. In a preferred embodiment, the extended cDNAs encode a full length protein sequence, which includes the protein coding sequences of SEQ ID NOs: 38-195.

EXAMPLE 27

General Method for Using 5′ ESTs to Clone and Sequence cDNAs which Include the Entire Coding Region and the Authentic 5′ End of the Corresponding mRNA

The following general method has been used to quickly and efficiently isolate extended cDNAs having the authentic 5′ ends of their corresponding mRNAs as well as

the full protein coding sequence and including sequence adjacent to the sequences of the 5′ ESTs used to obtain them. This method may be applied to obtain extended cDNAs for any 5′ EST in the NetGene™ database, including those 5′ ESTs encoding polypeptides belonging to secreted proteins. The method is summarized in figure 3.

1. Obtention of Extended cDNAs

a) First strand synthesis

The method takes advantage of the known 5′ sequence of the mRNA. A reverse transcription reaction is conducted on purified mRNA with a poly 14dT primer containing a 49 nucleotide sequence at its 5′ end allowing the addition of a known sequence at the end of the cDNA which corresponds to the 3′ end of the mRNA. For example, the primer may have the following sequence: 5′-ATC GTT GAG ACT CGT ACC AGC AGA GTC ACG AGA GAG ACT ACA CGG TAC TGG TTT TTT TTT TTT TTVN -3′ (SEQ ID NO:14). Those skilled in the art will appreciate that other sequences may also be added to the poly dT sequence and used to prime the first strand synthesis. Using this primer and a reverse transcriptase such as the Superscript II (Gibco BRL) or Rnase H Minus M-MLV (Promega) enzyme, a reverse transcript anchored at the 3′ polyA site of the RNAs is generated.

After removal of the mRNA hybridized to the first cDNA strand by alkaline hydrolysis, the products of the alkaline hydrolysis and the residual poly dT primer are eliminated with an exclusion column such as an AcA34 (Biosepra) matrix as explained in Example 11.

b) Second strand synthesis

A pair of nested primers on each end is designed based on the known 5′ sequence from the 5′ EST and the known 3′ end added by the poly dT primer used in the first strand synthesis. Softwares used to design primers are either based on GC content and melting temperatures of oligonucleotides, such as OSP (Illier and Green, *PCR Meth. Appl.* 1:124-128, 1991), or based on the octamer frequency disparity method (Griffais *et al.*, *Nucleic Acids Res.* **19**: 3887-3891, 1991) such as PC-Rare (http://bioinformatics.weizmann.ac.il/software/PC-Rare/doc/manuel.html).

Preferably, the nested primers at the 5′ end are separated from one another by four to nine bases. The 5′ primer sequences may be selected to have melting temperatures and specificities suitable for use in PCR.

Preferably, the nested primers at the 3′ end are separated from one another by four to nine bases. For example, the nested 3′ primers may have the following sequences: (5′- CCA GCA GAG TCA CGA GAG AGA CTA CAC GG -3′(SEQ ID NO:15), and 5′- CAC GAG AGA GAC TAC ACG GTA CTG G -3′ (SEQ ID NO:16). These primers were selected because they have melting temperatures and specificities compatible with their use in PCR. However, those skilled in the art will appreciate that other sequences may also be used as primers.

The first PCR run of 25 cycles is performed using the Advantage Tth Polymerase Mix (Clontech) and the outer primer from each of the nested pairs. A second 20 cycle PCR using the same enzyme and the inner primer from each of the nested pairs is then performed on 1/2500 of the first PCR product. Thereafter, the primers and nucleotides are removed.

<u>2. Sequencing of Full Length Extended cDNAs or Fragments Thereof</u>

Due to the lack of position constraints on the design of 5' nested primers compatible for PCR use using the OSP software, amplicons of two types are obtained. Preferably, the second 5' primer is located upstream of the translation initiation codon thus yielding a nested PCR product containing the whole coding sequence. Such a full length extended cDNA undergoes a direct cloning procedure as described in section a. However, in some cases, the second 5' primer is located downstream of the translation initiation codon, thereby yielding a PCR product containing only part of the ORF. Such incomplete PCR products are submitted to a modified procedure described in section b.

a) Nested PCR products containing complete ORFs

When the resulting nested PCR product contains the complete coding sequence, as predicted from the 5'EST sequence, it is cloned in an appropriate vector such as pED6dpc2, as described in section 3.

b) Nested PCR products containing incomplete ORFs

When the amplicon does not contain the complete coding sequence, intermediate steps are necessary to obtain both the complete coding sequence and a PCR product containing the full coding sequence. The complete coding sequence can be assembled from several partial sequences determined directly from different PCR products as described in the following section.

Once the full coding sequence has been completely determined, new primers compatible for PCR use are designed to obtain amplicons containing the whole coding region. However, in such cases, 3' primers compatible for PCR use are located inside the 3' UTR of the corresponding mRNA, thus yielding amplicons which lack part of this region, *i.e.* the polyA tract and sometimes the polyadenylation signal, as illustrated in figure 3. Such full length extended cDNAs are then cloned into an appropriate vector as described in section 3.

c) Sequencing extended cDNAs

Sequencing of extended cDNAs is performed using a Die Terminator approach with the AmpliTaq DNA polymerase FS kit available from Perkin Elmer.

In order to sequence PCR fragments, primer walking is performed using software such as OSP to choose primers and automated computer software such as ASMG (Sutton *et al., Genome Science Technol.* 1: 9-19, 1995) to construct contigs of walking sequences including the initial 5' tag using minimum overlaps of 32 nucleotides. Preferably, primer walking is performed until the sequences of full length cDNAs are obtained.

Completion of the sequencing of a given extended cDNA fragment is assessed as follows. Since sequences located after a polyA tract are difficult to determine precisely in the case of uncloned products, sequencing and primer walking processes for PCR products are interrupted when a polyA tract is identified in extended cDNAs obtained as described in case b. The sequence length is compared to the size of the nested PCR product obtained as described above. Due to the limited accuracy of the determination of the PCR product size by gel electrophoresis, a sequence is considered complete if the size of the obtained sequence is at least 70 % the size of the first nested PCR product. If the length of the sequence determined from the computer analysis is not at least 70% of the length of the nested PCR product, these PCR products are cloned and the sequence of the insertion is determined. When Northern blot data are available, the size of the mRNA detected for a given PCR

product is used to finally assess that the sequence is complete. Sequences which do not fulfill the above criteria are discarded and will undergo a new isolation procedure.

Sequence data of all extended cDNAs are then transferred to a proprietary database, where quality controls and validation steps are carried out as described in example 15.

3. Cloning of Full Length Extended cDNAs

The PCR product containing the full coding sequence is then cloned in an appropriate vector. For example, the extended cDNAs can be cloned into the expression vector pED6dpc2 (DiscoverEase, Genetics Institute, Cambridge, MA) as follows. pED6dpc2 vector DNA is prepared with blunt ends by performing an EcoRI digestion followed by a fill in reaction. The blunt ended vector is dephosphorylated. After removal of PCR primers and ethanol precipitation, the PCR product containing the full coding sequence or the extended cDNA obtained as described above is phosphorylated with a kinase subsequently removed by phenol-Sevag extraction and precipitation. The double stranded extended cDNA is then ligated to the vector and the resulting expression plasmid introduced into appropriate host cells.

Since the PCR products obtained as described above are blunt ended molecules that can be cloned in either direction, the orientation of several clones for each PCR product is determined. Then, 4 to 10 clones are ordered in microtiter plates and subjected to a PCR reaction using a first primer located in the vector close to the cloning site and a second primer located in the portion of the extended cDNA corresponding to the 3′ end of the mRNA. This second primer may be the antisense primer used in anchored PCR in the case of direct cloning (case a) or the antisense primer located inside the 3'UTR in the case of indirect cloning (case b). Clones in which the start codon of the extended cDNA is operably linked to the promoter in the vector so as to permit expression of the protein encoded by the extended cDNA are conserved and sequenced. In addition to the ends of cDNA inserts, approximately 50 bp of vector DNA on each side of the cDNA insert are also sequenced.

The cloned PCR products are then entirely sequenced according to the aforementioned procedure. In this case, contigation of long fragments is then performed on walking sequences that have already contigated for uncloned PCR products during

primer walking. Sequencing of cloned amplicons is complete when the resulting contigs include the whole coding region as well as overlapping sequences with vector DNA on both ends.

4. Computer analysis of Full Length Extended cDNA

Sequences of all full length extended cDNAs are then submitted to further analysis as described below. Before searching the extended full length cDNAs for sequences of interest, extended cDNAs which are not of interest (vector RNAs, transfer RNAs, ribosomal RNAs, mitochondrial RNAs, prokaryotic RNAs and fungal RNAs) are discarded using methods essentially similar to those described for 5'ESTs in Example 18.

a) Identification of structural features

Structural features, e.g. polyA tail and polyadenylation signal, of the sequences of full length extended cDNAs are subsequently determined as follows.

A polyA tail is defined as a homopolymeric stretch of at least 11 A with at most one alternative base within it. The polyA tail search is restricted to the last 100 nt of the sequence and limited to stretches of 11 consecutive A's because sequencing reactions are often not readable after such a polyA stretch. Stretches having more than 90% homology over 8 nucleotides are identified as polyA tails using BLAST2N.

To search for a polyadenylation signal, the polyA tail is clipped from the full-length sequence. The 50 bp preceding the polyA tail are first searched for the canonic polyadenylation AAUAAA signal and, if the canonic signal is not detected, for the alternative AUUAAA signal (Sheets *et al.*, *Nuc. Acids Res.* **18**: 5799-5805, 1990). If neither of these consensus polyadenylation signals is found, the canonic motif is searched again allowing one mismatch to account for possible sequencing errors. More than 85 % of identified polyadenylation signals of either type actually ends 10 to 30 bp from the polyA tail. Alternative AUUAAA signals represents approximately 15 % of the total number of identified polyadenylation signals.

b) Identification of functional features

Functional features, e.g. ORFs and signal sequences, of the sequences of full length extended cDNAs were subsequently determined as follows.

The 3 upper strand frames of extended cDNAs are searched for ORFs defined as the maximum length fragments beginning with a translation intiation codon and ending with a stop codon. ORFs encoding at least 20 amino acids are preferred.

Each found ORF is then scanned for the presence of a signal peptide in the first 50 amino-acids or, where appropriate, within shorter regions down to 20 amino acids or less in the ORF, using the matrix method of von Heijne (*Nuc. Acids Res.* **14**: 4683-4690, 1986), the disclosure of which is incorporated herein by reference as described in Example 22.

c) Homology to either nucleotidic or proteic sequences

Categorization of full-length sequences may be achieved using procedures essentially similar to those described for 5'ESTs in Example 24.

Extended cDNAs prepared as described above may be subsequently engineered to obtain nucleic acids which include desired portions of the extended cDNA using conventional techniques such as subcloning, PCR, or *in vitro* oligonucleotide synthesis. For example, nucleic acids which include only the full coding sequences (*i.e.* the sequences encoding the signal peptide and the mature protein remaining after the signal peptide is cleaved off) may be obtained using techniques known to those skilled in the art. Alternatively, conventional techniques may be applied to obtain nucleic acids which contain only the coding sequences for the mature protein remaining after the signal peptide is cleaved off or nucleic acids which contain only the coding sequences for the signal peptides.

Similarly, nucleic acids containing any other desired portion of the coding sequences for the secreted protein may be obtained. For example, the nucleic acid may contain at least 10 consecutive bases of an extended cDNA such as one of the extended cDNAs described below. In another embodiment, the nucleic acid may contain at least 15 consecutive bases of an extended cDNA such as one of the extended cDNAs described below. Alternatively, the nucleic acid may contain at least 20 consecutive bases of an extended cDNA such as one of the extended cDNAs described below. In another embodiment, the nucleic acid may contain at least 25 consecutive bases of an extended cDNA such as one of the extended cDNAs described below. In yet another embodiment, the nucleic acid may contain at least 40

consecutive bases of an extended cDNA such as one of the extended cDNAs described below.

Once an extended cDNA has been obtained, it can be sequenced to determine the amino acid sequence it encodes. Once the encoded amino acid sequence has been determined, one can create and identify any of the many conceivable cDNAs that will encode that protein by simply using the degeneracy of the genetic code. For example, allelic variants or other homologous nucleic acids can be identified as described below. Alternatively, nucleic acids encoding the desired amino acid sequence can be synthesized *in vitro*.

In a preferred embodiment, the coding sequence may be selected using the known codon or codon pair preferences for the host organism in which the cDNA is to be expressed.

The extended cDNAs derived from the 5' ESTS of the present invention were obtained as described in Example 28 below.

EXAMPLE 28

Characterization of cloned extended cDNAs obtained using 5' ESTs

The procedure described in Example 27 above was used to obtain the extended cDNAs derived from the 5' ESTs of the present invention in a variety of tissues. The following list provides a few examples of thus obtained extended cDNAs.

Using this approach, the full length cDNA of SEQ ID NO:17 (internal identification number 48-19-3-G1-FL1) was obtained. This cDNA falls into the "EST-ext" category described above and encodes the signal peptide MKKVLLLITAILAVAVG (SEQ ID NO: 18) having a von Heijne score of 8.2.

The full length cDNA of SEQ ID NO:19 (internal identification number 58-34-2-E7-FL2) was also obtained using this procedure. This cDNA falls into the "EST-ext" category described above and encodes the signal peptide MWWFQQGLSFLPSALVIWTSA (SEQ ID NO:20) having a von Heijne score of 5.5.

Another full length cDNA obtained using the procedure described above has the sequence of SEQ ID NO:21 (internal identification number 51-27-1-E8-FL1). This cDNA, falls into the "EST-ext" category described above and encodes the signal peptide MVLTTLPSANSANSPVNMPTTGPNSLSYASSALSPCLT (SEQ ID NO:22) having a von Heijne score of 5.9.

The above procedure was also used to obtain a full length cDNA having the sequence of SEQ ID NO:23 (internal identification number 76-4-1-G5-FL1). This cDNA falls into the "EST-ext" category described above and encodes the signal peptide ILSTVTALTFAXA (SEQ ID NO:24) having a von Heijne score of 5.5.

The full length cDNA of SEQ ID NO:25 (internal identification number 51-3-3-B10-FL3) was also obtained using this procedure. This cDNA falls into the "new" category described above and encodes a signal peptide LVLTLCTLPLAVA (SEQ ID NO:26) having a von Heijne score of 10.1.

The full length cDNA of SEQ ID NO:27 (internal identification number 58-35-2-F10-FL2) was also obtained using this procedure. This cDNA falls into the "new" category described above and encodes a signal peptide LWLLFFLVTAIHA (SEQ ID NO:28) having a von Heijne score of 10.7.

Bacterial clones containing plasmids containing the full length cDNAs described above are presently stored in the inventor's laboratories under the internal identification numbers provided above. The inserts may be recovered from the stored materials by growing an aliquot of the appropriate bacterial clone in the appropriate medium. The plasmid DNA can then be isolated using plasmid isolation procedures familiar to those skilled in the art such as alkaline lysis minipreps or large scale alkaline lysis plasmid isolation procedures. If desired the plasmid DNA may be further enriched by centrifugation on a cesium chloride gradient, size exclusion chromatography, or anion exchange chromatography. The plasmid DNA obtained using these procedures may then be manipulated using standard cloning techniques familiar to those skilled in the art. Alternatively, a PCR can be done with primers designed at both ends of the cDNA insertion. The PCR product which corresponds to the cDNA can then be manipulated using standard cloning techniques familiar to those skilled in the art.

The polypeptides encoded by the extended cDNAs may be screened for the presence of known structural or functional motifs or for the presence of signatures, small amino acid sequences which are well conserved amongst the members of a protein family. The conserved regions have been used to derive consensus patterns or matrices included in the PROSITE data bank, in particular in the file prosite.dat (Release 13.0 of November 1995, located at http://expasy.hcuge.ch/sprot/prosite.html. Prosite_convert and prosite_scan

programs (http://ulrec3.unil.ch/ftpserveur/prosite_scan) may be used to find signatures on the extended cDNAs.

For each pattern obtained with the prosite_convert program from the prosite.dat file, the accuracy of the detection on a new protein sequence may be assessed by evaluating the frequency of irrelevant hits on the population of human secreted proteins included in the data bank SWISSPROT. The ratio between the number of hits on shuffled proteins (with a window size of 20 amino acids) and the number of hits on native (unshuffled) proteins may be used as an index. Every pattern for which the ratio is greater than 20% (one hit on shuffled proteins for 5 hits on native proteins) may be skipped during the search with prosite_scan. The program used to shuffle protein sequences (db_shuffled) and the program used to determine the statistics for each pattern in the protein data banks (prosite_statistics) are available on the ftp site http://ulrec3.unil.ch/ftpserveur/prosite_scan.

In addition to PCR based methods for obtaining extended cDNAs, traditional hybridization based methods may also be employed. These methods may also be used to obtain the genomic DNAs which encode the mRNAs from which the 5′ ESTs were derived, mRNAs corresponding to the extended cDNAs, or nucleic acids which are homologous to extended cDNAs or 5′ ESTs. Example 29 below provides examples of such methods.

EXAMPLE 29

Methods for Obtaining cDNAs which include the Entire Coding Region and the Authentic 5'End of the Corresponding mRNA

A full length cDNA library can be made using the strategies described in Examples 13, 14, 15, and 16 above by replacing the random nonamer used in Example 14 with an oligo-dT primer. For instance, the oligonucleotide of SEQ ID NO:14 may be used.

Alternatively, a cDNA library or genomic DNA library may be obtained from a commercial source or made using techniques familiar to those skilled in the art. Such cDNA or genomic DNA librairies may be used to isolate extended cDNAs obtained from 5′ EST or nucleic acids homologous to extended cDNAs or 5′ EST as follows. The cDNA library or genomic DNA library is hybridized to a detectable probe comprising at least 10 consecutive nucleotides from the 5′ EST or extended cDNA using conventional techniques. Preferably,

the probe comprises at least 12, 15, or 17 consecutive nucleotides from the 5′ EST or extended cDNA. More preferably, the probe comprises at least 20 to 30 consecutive nucleotides from the 5′ EST or extended cDNA. In some embodiments, the probe comprises more than 30 nucleotides from the 5′ EST or extended cDNA.

Techniques for identifying cDNA clones in a cDNA library which hybridize to a given probe sequence are disclosed in Sambrook *et al., Molecular Cloning: A Laboratory Manual 2d Ed.*, Cold Spring Harbor Laboratory Press, 1989, the disclosure of which is incorporated herein by reference. The same techniques may be used to isolate genomic DNAs.

Briefly, cDNA or genomic DNA clones which hybridize to the detectable probe are identified and isolated for further manipulation as follows. A probe comprising at least 10 consecutive nucleotides from the 5′ EST or extended cDNA is labeled with a detectable label such as a radioisotope or a fluorescent molecule. Preferably, the probe comprises at least 12, 15, or 17 consecutive nucleotides from the 5′ EST or extended cDNA. More preferably, the probe comprises 20 to 30 consecutive nucleotides from the 5′ EST or extended cDNA. In some embodiments, the probe comprises more than 30 nucleotides from the 5′ EST or extended cDNA.

Techniques for labeling the probe are well known and include phosphorylation with polynucleotide kinase, nick translation, *in vitro* transcription, and non radioactive techniques. The cDNAs or genomic DNAs in the library are transferred to a nitrocellulose or nylon filter and denatured. After blocking of non specific sites, the filter is incubated with the labeled probe for an amount of time sufficient to allow binding of the probe to cDNAs or genomic DNAs containing a sequence capable of hybridizing thereto.

By varying the stringency of the hybridization conditions used to identify extended cDNAs or genomic DNAs which hybridize to the detectable probe, extended cDNAS having different levels of homology to the probe can be identified and isolated as described below.

1. Identification of Extended cDNA or Genomic cDNA Sequences Having a High Degree of Homology to the Labeled Probe

To identify extended cDNAs or genomic DNAs having a high degree of homology to the probe sequence, the melting temperature of the probe may be calculated using the following formulas:

For probes between 14 and 70 nucleotides in length the melting temperature (Tm) is calculated using the formula: Tm=81.5+16.6(log [Na+])+0.41(fraction G+C)-(600/N) where N is the length of the probe.

If the hybridization is carried out in a solution containing formamide, the melting temperature may be calculated using the equation Tm=81.5+16.6(log [Na+])+0.41(fraction G+C)-(0.63% formamide)-(600/N) where N is the length of the probe.

Prehybridization may be carried out in 6X SSC, 5X Denhardt's reagent, 0.5% SDS, 100 μg denatured fragmented salmon sperm DNA or 6X SSC, 5X Denhardt's reagent, 0.5% SDS, 100 μg denatured fragmented salmon sperm DNA, 50% formamide. The formulas for SSC and Denhardt's solutions are listed in Sambrook *et al.*, *supra*.

Hybridization is conducted by adding the detectable probe to the prehybridization solutions listed above. Where the probe comprises double stranded DNA, it is denatured before addition to the hybridization solution. The filter is contacted with the hybridization solution for a sufficient period of time to allow the probe to hybridize to extended cDNAs or genomic DNAs containing sequences complementary thereto or homologous thereto. For probes over 200 nucleotides in length, the hybridization may be carried out at 15-25°C below the Tm. For shorter probes, such as oligonucleotide probes, the hybridization may be conducted at 15-25°C below the Tm. Preferably, for hybridizations in 6X SSC, the hybridization is conducted at approximately 68°C. Preferably, for hybridizations in 50% formamide containing solutions, the hybridization is conducted at approximately 42°C.

All of the foregoing hybridizations would be considered to be under "stringent" conditions.

Following hybridization, the filter is washed in 2X SSC, 0.1% SDS at room temperature for 15 minutes. The filter is then washed with 0.1X SSC, 0.5% SDS at room temperature for 30 minutes to 1 hour. Thereafter, the solution is washed at the hybridization

temperature in 0.1X SSC, 0.5% SDS. A final wash is conducted in 0.1X SSC at room temperature.

Extended cDNAs, nucleic acids homologous to extended cDNAs or 5′ ESTs, or genomic DNAs which have hybridized to the probe are identified by autoradiography or other conventional techniques.

2. Obtention of Extended cDNA or Genomic cDNA Sequences Having Lower Degrees of Homology to the Labeled Probe

The above procedure may be modified to identify extended cDNAs, nucleic acids homologous to extended cDNAs, or genomic DNAs having decreasing levels of homology to the probe sequence. For example, to obtain extended cDNAs, nucleic acids homologous to extended cDNAs, or genomic DNAs of decreasing homology to the detectable probe, less stringent conditions may be used. For example, the hybridization temperature may be decreased in increments of 5°C from 68°C to 42°C in a hybridization buffer having a sodium concentration of approximately 1M. Following hybridization, the filter may be washed with 2X SSC, 0.5% SDS at the temperature of hybridization. These conditions are considered to be "moderate" conditions above 50°C and "low" conditions below 50°C.

Alternatively, the hybridization may be carried out in buffers, such as 6X SSC, containing formamide at a temperature of 42°C. In this case, the concentration of formamide in the hybridization buffer may be reduced in 5% increments from 50% to 0% to identify clones having decreasing levels of homology to the probe. Following hybridization, the filter may be washed with 6X SSC, 0.5% SDS at 50°C. These conditions are considered to be "moderate" conditions above 25% formamide and "low" conditions below 25% formamide.

Extended cDNAs, nucleic acids homologous to extended cDNAs, or genomic DNAs which have hybridized to the probe are identified by autoradiography.

3. Determination of the Degree of Homology Between the Obtained Extended cDNAs and the Labeled Probe

If it is desired to obtain nucleic acids homologous to extended cDNAs, such as allelic variants thereof or nucleic acids encoding proteins related to the proteins encoded by the extended cDNAs, the level of homology between the hybridized nucleic acid and the

extended cDNA or 5′ EST used as the probe may be further determined using BLAST2N; parameters may be adapted depending on the sequence length and degree of homology studied. To determine the level of homology between the hybridized nucleic acid and the extended cDNA or 5′EST from which the probe was derived, the nucleotide sequences of the hybridized nucleic acid and the extended cDNA or 5′EST from which the probe was derived are compared. For example, using the above methods, nucleic acids having at least 95% nucleic acid homology to the extended cDNA or 5′EST from which the probe was derived may be obtained and identified. Similarly, by using progressively less stringent hybridization conditions one can obtain and identify nucleic acids having at least 90%, at least 85%, at least 80% or at least 75% homology to the extended cDNA or 5′EST from which the probe was derived.

To determine whether a clone encodes a protein having a given amount of homology to the protein encoded by the extended cDNA or 5′ EST, the amino acid sequence encoded by the extended cDNA or 5′ EST is compared to the amino acid sequence encoded by the hybridizing nucleic acid. Homology is determined to exist when an amino acid sequence in the extended cDNA or 5′ EST is closely related to an amino acid sequence in the hybridizing nucleic acid. A sequence is closely related when it is identical to that of the extended cDNA or 5′ EST or when it contains one or more amino acid substitutions therein in which amino acids having similar characteristics have been substituted for one another. Using the above methods and algorithms such as FASTA with parameters depending on the sequence length and degree of homology studied, one can obtain nucleic acids encoding proteins having at least 95%, at least 90%, at least 85%, at least 80% or at least 75% homology to the proteins encoded by the extended cDNA or 5′EST from which the probe was derived.

In addition to the above described methods, other protocols are available to obtain extended cDNAs using 5' ESTs as outlined in the following paragraphs.

Extended cDNAs may be prepared by obtaining mRNA from the tissue, cell, or organism of interest using mRNA preparation procedures utilizing polyA selection procedures or other techniques known to those skilled in the art. A first primer capable of hybridizing to the polyA tail of the mRNA is hybridized to the mRNA and a reverse transcription reaction is performed to generate a first cDNA strand.

The first cDNA strand is hybridized to a second primer containing at least 10 consecutive nucleotides of the sequences of SEQ ID NOs 38-195. Preferably, the primer comprises at least 12, 15, or 17 consecutive nucleotides from the sequences of SEQ ID NOs 38-195. More preferably, the primer comprises 20 to 30 consecutive nucleotides from the sequences of SEQ ID NOs 38-195. In some embodiments, the primer comprises more than 30 nucleotides from the sequences of SEQ ID NOs 38-195. If it is desired to obtain extended cDNAs containing the full protein coding sequence, including the authentic translation initiation site, the second primer used contains sequences located upstream of the translation initiation site. The second primer is extended to generate a second cDNA strand complementary to the first cDNA strand. Alternatively, RT-PCR may be performed as described above using primers from both ends of the cDNA to be obtained.

Extended cDNAs containing 5′ fragments of the mRNA may be prepared by hybridizing an mRNA comprising the sequence of the 5'EST for which an extended cDNA is desired with a primer comprising at least 10 consecutive nucleotides of the sequences complementary to the 5'EST and reverse transcribing the hybridized primer to make a first cDNA strand from the mRNAs. Preferably, the primer comprises at least 12, 15, or 17 consecutive nucleotides from the 5'EST. More preferably, the primer comprises 20 to 30 consecutive nucleotides from the 5'EST.

Thereafter, a second cDNA strand complementary to the first cDNA strand is synthesized. The second cDNA strand may be made by hybridizing a primer complementary to sequences in the first cDNA strand to the first cDNA strand and extending the primer to generate the second cDNA strand.

The double stranded extended cDNAs made using the methods described above are isolated and cloned. The extended cDNAs may be cloned into vectors such as plasmids or viral vectors capable of replicating in an appropriate host cell. For example, the host cell may be a bacterial, mammalian, avian, or insect cell.

Techniques for isolating mRNA, reverse transcribing a primer hybridized to mRNA to generate a first cDNA strand, extending a primer to make a second cDNA strand complementary to the first cDNA strand, isolating the double stranded cDNA and cloning the double stranded cDNA are well known to those skilled in the art and are described *in Current Protocols in Molecular Biology*, John Wiley and Sons, Inc. 1997 and Sambrook *et al.*,

Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, 1989, the entire disclosures of which are incorporated herein by reference.

Alternatively, procedures such as the one described in Example 29 may be used for obtaining full length cDNAs or extended cDNAs. In this approach, full length or extended cDNAs are prepared from mRNA and cloned into double stranded phagemids as follows. The cDNA library in the double stranded phagemids is then rendered single stranded by treatment with an endonuclease, such as the Gene II product of the phage F1, and an exonuclease (Chang *et al.*, *Gene* **127**:95-8, 1993). A biotinylated oligonucleotide comprising the sequence of a 5′ EST, or a fragment containing at least 10 nucleotides thereof, is hybridized to the single stranded phagemids. Preferably, the fragment comprises at least 12, 15, or 17 consecutive nucleotides from the 5′ EST. More preferably, the fragment comprises 20-30 consecutive nucleotides from the 5′ EST. In some procedures, the fragment may comprise more than 30 consecutive nucleotides from the 5′ EST.

Hybrids between the biotinylated oligonucleotide and phagemids having inserts containing the 5′ EST sequence are isolated by incubating the hybrids with streptavidin coated paramagnetic beads and retrieving the beads with a magnet (Fry *et al.*, *Biotechniques*, **13**: 124-131, 1992). Therafter, the resulting phagemids containing the 5′ EST sequence are released from the beads and converted into double stranded DNA using a primer specific for the 5′ EST sequence. Alternatively, protocoles such as the Gene Trapper kit (Gibco BRL) may be used. The resulting double stranded DNA is transformed into bacteria. Extended cDNAs containing the 5′ EST sequence are identified by colony PCR or colony hybridization.

Using any of the above described methods in section III, a plurality of extended cDNAs containing full length protein coding sequences or sequences encoding only the mature protein remaining after the signal peptide is cleaved off may be provided as cDNA libraries for subsequent evaluation of the encoded proteins or use in diagnostic assays as described below.

IV. Expression of Proteins Encoded by Extended cDNAs Isolated Using 5′ ESTs

Extended cDNAs containing the full protein coding sequences of their corresponding mRNAs or portions thereof, such as cDNAs encoding the mature protein, may be used to

express the encoded secreted proteins or portions thereof as described in Example 30 below. If desired, the extended cDNAs may contain the sequences encoding the signal peptide to facilitate secretion of the expressed protein. It will be appreciated that a plurality of extended cDNAs containing the full protein coding sequences or portions thereof may be simultaneously cloned into expression vectors to create an expression library for analysis of the encoded proteins as described below.

EXAMPLE 30

Expression of the Proteins Encoded by the Genes Corresponding to 5'ESTS or Portions Thereof

To express the proteins encoded by the genes corresponding to 5′ ESTs (or portions thereof), full length cDNAs containing the entire protein coding region or extended cDNAs containing sequences adjacent to the 5′ ESTs (or portions thereof) are obtained as described in Examples 27-29 and cloned into a suitable expression vector. If desired, the nucleic acids may contain the sequences encoding the signal peptide to facilitate secretion of the expressed protein. The nucleic acids inserted into the expression vectors may also contain sequences upstream of the sequences encoding the signal peptide, such as sequences which regulate expression levels or sequences which confer tissue specific expression.

The nucleic acid encoding the protein or polypeptide to be expressed is operably linked to a promoter in an expression vector using conventional cloning technology. The expression vector may be any of the mammalian, yeast, insect or bacterial expression systems known in the art. Commercially available vectors and expression systems are available from a variety of suppliers including Genetics Institute (Cambridge, MA), Stratagene (La Jolla, California), Promega (Madison, Wisconsin), and Invitrogen (San Diego, California). If desired, to enhance expression and facilitate proper protein folding, the codon context and codon pairing of the sequence may be optimized for the particular expression organism in which the expression vector is introduced, as explained by Hatfield, *et al.*, U.S. Patent No. 5,082,767, incorporated herein by this reference.

The cDNA cloned into the expression vector may encode the entire protein (*i.e.* the signal peptide and the mature protein), the mature protein (*i.e.* the protein created by cleaving the signal peptide off), only the signal peptide or any other portion thereof.

The following is provided as one exemplary method to express the proteins encoded by the extended cDNAs corresponding to the 5′ ESTs or the nucleic acids described above. First, the methionine initiation codon for the gene and the polyA signal of the gene are identified. If the nucleic acid encoding the polypeptide to be expressed lacks a methionine to serve as the initiation site, an initiating methionine can be introduced next to the first codon of the nucleic acid using conventional techniques. Similarly, if the extended cDNA lacks a polyA signal, this sequence can be added to the construct by, for example, splicing out the polyA signal from pSG5 (Stratagene) using BglII and SalI restriction endonuclease enzymes and incorporating it into the mammalian expression vector pXT1 (Stratagene). pXT1 contains the LTRs and a portion of the *gag* gene from Moloney Murine Leukemia Virus. The position of the LTRs in the construct allow efficient stable transfection. The vector includes the Herpes Simplex thymidine kinase promoter and the selectable neomycin gene. The extended cDNA or portion thereof encoding the polypeptide to be expressed is obtained by PCR from the bacterial vector using oligonucleotide primers complementary to the extended cDNA or portion thereof and containing restriction endonuclease sequences for Pst I incorporated into the 5′primer and BglII at the 5′ end of the corresponding cDNA 3′ primer, taking care to ensure that the extended cDNA is positioned with the poly A signal. The purified fragment obtained from the resulting PCR reaction is digested with PstI, blunt ended with an exonuclease, digested with Bgl II, purified and ligated to pXT1 containing a poly A signal and prepared for this ligation (blunt/BglII).

The ligated product is transfected into mouse NIH 3T3 cells using Lipofectin (Life Technologies, Inc., Grand Island, New York) under conditions outlined in the product specification. Positive transfectants are selected after growing the transfected cells in 600 µg/ml G418 (Sigma, St. Louis, Missouri). Preferably the expressed protein is released into the culture medium, thereby facilitating purification.

Alternatively, the extended cDNAs may be cloned into pED6dpc2 as described above. The resulting pED6dpc2 constructs may be transfected into a suitable host cell, such as COS 1 cells. Methotrexate resistant cells are selected and expanded. Preferably, the protein expressed from the extended cDNA is released into the culture medium thereby facilitating purification.

Proteins in the culture medium are separated by gel electrophoresis. If desired, the proteins may be ammonium sulfate precipitated or separated based on size or charge prior to electrophoresis.

As a control, the expression vector lacking a cDNA insert is introduced into host cells or organisms and the proteins in the medium are harvested. The secreted proteins present in the medium are detected using techniques familiar to those skilled in the art such as Coomassie blue or silver staining or using antibodies against the protein encoded by the extended cDNA.

Antibodies capable of specifically recognizing the protein of interest may be generated using synthetic 15-mer peptides having a sequence encoded by the appropriate 5′ EST, extended cDNA, or portion thereof. The synthetic peptides are injected into mice to generate antibody to the polypeptide encoded by the 5′ EST, extended cDNA, or portion thereof.

Secreted proteins from the host cells or organisms containing an expression vector which contains the extended cDNA derived from a 5′ EST or a portion thereof are compared to those from the control cells or organism. The presence of a band in the medium from the cells containing the expression vector which is absent in the medium from the control cells indicates that the extended cDNA encodes a secreted protein. Generally, the band corresponding to the protein encoded by the extended cDNA will have a mobility near that expected based on the number of amino acids in the open reading frame of the extended cDNA. However, the band may have a mobility different than that expected as a result of modifications such as glycosylation, ubiquitination, or enzymatic cleavage.

Alternatively, if the protein expressed from the above expression vectors does not contain sequences directing its secretion, the proteins expressed from host cells containing an expression vector with an insert encoding a secreted protein or portion thereof can be compared to the proteins expressed in control host cells containing the expression vector without an insert. The presence of a band in samples from cells containing the expression vector with an insert which is absent in samples from cells containing the expression vector without an insert indicates that the desired protein or portion thereof is being expressed. Generally, the band will have the mobility expected for the secreted protein or portion thereof. However, the band may have a mobility different than that expected as a result of modifications such as glycosylation, ubiquitination, or enzymatic cleavage.

The protein encoded by the extended cDNA may be purified using standard immunochromatography techniques. In such procedures, a solution containing the secreted protein, such as the culture medium or a cell extract, is applied to a column having antibodies against the secreted protein attached to the chromatography matrix. The secreted protein is allowed to bind the immunochromatography column. Thereafter, the column is washed to remove non-specifically bound proteins. The specifically bound secreted protein is then released from the column and recovered using standard techniques.

If antibody production is not possible, the extended cDNA sequence or portion thereof may be incorporated into expression vectors designed for use in purification schemes employing chimeric polypeptides. In such strategies, the coding sequence of the extended cDNA or portion thereof is inserted in frame with the gene encoding the other half of the chimera. The other half of the chimera may be β-globin or a nickel binding polypeptide. A chromatography matrix having antibody to β-globin or nickel attached thereto is then used to purify the chimeric protein. Protease cleavage sites may be engineered between the β-globin gene or the nickel binding polypeptide and the extended cDNA or portion thereof. Thus, the two polypeptides of the chimera may be separated from one another by protease digestion.

One useful expression vector for generating β-globin chimerics is pSG5 (Stratagene), which encodes rabbit β-globin. Intron II of the rabbit β-globin gene facilitates splicing of the expressed transcript, and the polyadenylation signal incorporated into the construct increases the level of expression. These techniques as described are well known to those skilled in the art of molecular biology. Standard methods are published in methods texts such as Davis *et al.*., (*Basic Methods in Molecular Biology*, Davis, Dibner, and Battey, ed., Elsevier Press, NY, 1986) and many of the methods are available from Stratagene, Life Technologies, Inc., or Promega. Polypeptide may additionally be produced from the construct using *in vitro* translation systems such as the *In vitro* Express™ Translation Kit (Stratagene).

Following expression and purification of the secreted proteins encoded by the 5′ ESTs, extended cDNAs, or fragments thereof, the purified proteins may be tested for the ability to bind to the surface of various cell types as described in Example 31 below. It will be appreciated that a plurality of proteins expressed from these cDNAs may be included in a

panel of proteins to be simultaneously evaluated for the activities specifically described below, as well as other biological roles for which assays for determining activity are available.

EXAMPLE 31

Analysis of Secreted Proteins to Determine Whether they Bind to the Cell Surface

The proteins encoded by the 5′ ESTs, extended cDNAs, or fragments thereof are cloned into expression vectors such as those described in Example 30. The proteins are purified by size, charge, immunochromatography or other techniques familiar to those skilled in the art. Following purification, the proteins are labeled using techniques known to those skilled in the art. The labeled proteins are incubated with cells or cell lines derived from a variety of organs or tissues to allow the proteins to bind to any receptor present on the cell surface. Following the incubation, the cells are washed to remove non-specifically bound protein. The labeled proteins are detected by autoradiography. Alternatively, unlabeled proteins may be incubated with the cells and detected with antibodies having a detectable label, such as a fluorescent molecule, attached thereto.

Specificity of cell surface binding may be analyzed by conducting a competition analysis in which various amounts of unlabeled protein are incubated along with the labeled protein. The amount of labeled protein bound to the cell surface decreases as the amount of competitive unlabeled protein increases. As a control, various amounts of an unlabeled protein unrelated to the labeled protein is included in some binding reactions. The amount of labeled protein bound to the cell surface does not decrease in binding reactions containing increasing amounts of unrelated unlabeled protein, indicating that the protein encoded by the cDNA binds specifically to the cell surface.

As discussed above, secreted proteins have been shown to have a number of important physiological effects and, consequently, represent a valuable therapeutic resource. The secreted proteins encoded by the extended cDNAs or portions thereof made according to Examples 27-29 may be evaluated to determine their physiological activities as described below.

EXAMPLE 32

Assaying the Proteins Expressed from Extended cDNAs or Portions Thereof for Cytokine, Cell Proliferation or Cell Differentiation Activity

As discussed above, secreted proteins may act as cytokines or may affect cellular proliferation or differentiation. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein encoded by the extended cDNAs is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7c and CMK. The proteins encoded by the above extended cDNAs or portions thereof may be evaluated for their ability to regulate T cell or thymocyte proliferation in assays such as those described above or in the following references, which are incorporated herein by reference: *Current Protocols in Immunology*, Ed. by Coligan *et al.*., Greene Publishing Associates and Wiley-Interscience; Takai *et al. J. Immunol.* **137**:3494-3500, 1986., Bertagnolli *et al.*, J. *Immunol.* **145**:1706-1712, 1990., Bertagnolli *et al.*, *Cell. Immunol.* **133**:327-341, 1991; Bertagnolli, *et al.*, *J. Immunol.* **149**:3778-3783, 1992; Bowman *et al.*, *J. Immunol.* **152**:1756-1761, 1994.

In addition, numerous assays for cytokine production and/or the proliferation of spleen cells, lymph node cells and thymocytes are known. These include the techniques disclosed in *Current Protocols in Immunology, supra* **1**:3.12.1-3.12.14; and Schreiber In *Current Protocols in Immunology, supra* 1 : 6.8.1-6.8.8.

The proteins encoded by the cDNAs may also be assayed for the ability to regulate the proliferation and differentiation of hematopoietic or lymphopoietic cells. Many assays for such activity are familiar to those skilled in the art, including the assays in the following references, which are incorporated herein by reference: Bottomly *et al.*, In *Current Protocols in Immunology.*, *supra.* 1 : 6.3.1-6.3.12,; deVries *et al.*, *J. Exp. Med.* **173**:1205-1211, 1991; Moreau *et al.*, *Nature* **36**:690-692, 1988; Greenberger *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **80**:2931-2938, 1983; Nordan, R., In *Current Protocols in Immunology.*, *supra.* 1 : 6.6.1-6.6.5; Smith et al., *Proc. Natl. Acad. Sci. U.S.A.* **83**:1857-1861, 1986; Bennett *et al.*, in

Current Protocols in Immunology supra **1** : 6.15.1; Ciarletta *et al.*, In *Current Protocols in Immunology. supra* **1** : 6.13.1.

The proteins encoded by the cDNAs may also be assayed for their ability to regulate T-cell responses to antigens. Many assays for such activity are familiar to those skilled in the art, including the assays described in the following references, which are incorporated herein by reference: Chapter 3 (*In Vitro* Assays for Mouse Lymphocyte Function), Chapter 6 (Cytokines and Their Cellular Receptors) and Chapter 7, (Immunologic Studies in Humans) in *Current Protocols in Immunology supra*; Weinberger *et al.*, *Proc. Natl. Acad. Sci. USA* 77:6091-6095, 1980; Weinberger *et al.*, *Eur. J. Immun.* **11**:405-411, 1981; Takai *et al.*, *J. Immunol.* **137**:3494-3500, 1986; Takai *et al.*, *J. Immunol.* **140**:508-512, 1988.

Those proteins which exhibit cytokine, cell proliferation, or cell differentiation activity may then be formulated as pharmaceuticals and used to treat clinical conditions in which induction of cell proliferation or differentiation is beneficial. Alternatively, as described in more detail below, genes encoding these proteins or nucleic acids regulating the expression of these proteins may be introduced into appropriate host cells to increase or decrease the expression of the proteins as desired.

EXAMPLE 33

Assaying the Proteins Expressed from Extended cDNAs or Portions Thereof for Activity as Immune System Regulators

The proteins encoded by the cDNAs may also be evaluated for their effects as immune regulators. For example, the proteins may be evaluated for their activity to influence thymocyte or splenocyte cytotoxicity. Numerous assays for such activity are familiar to those skilled in the art including the assays described in the following references, which are incorporated herein by reference: Chapter 3 (*In Vitro* Assays for Mouse Lymphocyte Function 3.1-3.19) and Chapter 7 (Immunologic studies in Humans) in *Current Protocols in Immunology* , Coligan *et al.*, Eds,. Greene Publishing Associates and Wiley-Interscience; Herrmann *et al.*, *Proc. Natl. Acad. Sci. USA* **78**:2488-2492, 1981; Herrmann *et al.*, *J. Immunol.* **128**:1968-1974, 1982; Handa *et al.*, *J. Immunol.* **135**:1564-1572, 1985; Takai *et al.*, *J. Immunol.* 137:3494-3500, 1986; Takai *et al.*, *J. Immunol.* **140**:508-512, 1988;

Bowman *et al.*, *J. Virology* **61**:1992-1998; Bertagnolli *et al.*, *Cell. Immunol.* **133**:327-341, 1991; Brown *et al.*, *J. Immunol.* **153**:3079-3092, 1994.

The proteins encoded by the cDNAs may also be evaluated for their effects on T-cell dependent immunoglobulin responses and isotype switching. Numerous assays for such activity are familiar to those skilled in the art, including the assays disclosed in the following references, which are incorporated herein by reference: Maliszewski, *J. Immunol.* **144**:3028-3033, 1990; Mond *et al.* in *Current Protocols in Immunology*, 1 : 3.8.1-3.8.16, *supra.*

The proteins encoded by the cDNAs may also be evaluated for their effect on immune effector cells, including their effect on Th1 cells and cytotoxic lymphocytes. Numerous assays for such activity are familiar to those skilled in the art, including the assays disclosed in the following references, which are incorporated herein by reference: Chapter 3 (*In Vitro* Assays for Mouse Lymphocyte Function 3.1-3.19) and Chapter 7 (Immunologic Studies in Humans) in *Current Protocols in Immunology, supra*; Takai *et al., J. Immunol.* **137**:3494-3500, 1986; Takai *et al., J. Immunol.* **140**:508-512, 1988; Bertagnolli *et al., J. Immunol.* **149**:3778-3783, 1992.

The proteins encoded by the cDNAs may also be evaluated for their effect on dendritic cell mediated activation of naive T-cells. Numerous assays for such activity are familiar to those skilled in the art, including the assays disclosed in the following references, which are incorporated herein by reference: Guery *et al., J. Immunol.* **134**:536-544, 1995; Inaba *et al., J. Exp. Med.* **173**:549-559, 1991; Macatonia *et al., J. Immunol.* **154**:5071-5079, 1995; Porgador *et al.J. Exp. Med* **182**:255-260, 1995; Nair *et al., J. Virol.* **67**:4062-4069, 1993; Huang *et al., Science* **264**:961-965, 1994; Macatonia *et al.J. Exp. Med* **169**:1255-1264, 1989; Bhardwaj *et al., Journal of Clinical Investigation* **94**:797-807, 1994; and Inaba *et al., J. Exp. Med* **172**:631-640, 1990.

The proteins encoded by the cDNAs may also be evaluated for their influence on the lifetime of lymphocytes. Numerous assays for such activity are familiar to those skilled in the art, including the assays disclosed in the following references, which are incorporated herein by reference: Darzynkiewicz *et al., Cytometry* **13**:795-808, 1992; Gorczyca *et al., Leukemia* **7**:659-670, 1993; Gorczyca *et al., Cancer Res.* **53**:1945-1951, 1993; Itoh *et al., Cell* **66**:233-243, 1991; Zacharchuk, *J. Immunol.* **145**:4037-4045, 1990; Zamai *et al., Cytometry* **14**:891-897, 1993; Gorczyca *et al., Int. J. Oncol.* **1**:639-648, 1992.

The proteins encoded by the cDNAs may also be evaluated for their influence on early steps of T-cell commitment and development. Numerous assays for such activity are familiar to those skilled in the art, including without limitation the assays disclosed in the following references, which are incorporated herein by references: Antica *et al.*, *Blood* **84**:111-117, 1994; Fine *et al.*, *Cell. Immunol.* **155**:111-122, 1994; Galy *et al.*, *Blood* **85**:2770-2778, 1995; Toki *et al.*, *Proc. Nat. Acad Sci. USA* **88**:7548-7551, 1991.

Those proteins which exhibit activity as immune system regulators activity may then be formulated as pharmaceuticals and used to treat clinical conditions in which regulation of immune activity is beneficial. For example, the protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases caused by viral, bacterial, fungal or other infection may be treatable using a protein encoded by extended cDNAs derived from the 5' ESTs of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, Leishmania spp., plamodium and various fungal infections such as candidiasis. Of course, in this regard, a protein encoded by extended cDNAs derived from the 5' ESTs of the present invention may also be useful where a boost to the immune system generally may be desirable, *i.e.*, in the treatment of cancer.

Alternatively, proteins encoded by extended cDNAs derived from the 5' ESTs of the present invention may be used in treatment of autoimmune disorders including, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein encoded by extended cDNAs derived from the 5' ESTs of the present invention may also to be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein encoded by extended cDNAs derived from the 5' ESTs of the present invention.

Using the proteins of the invention it may also be possible to regulate immune responses either up or down.

Down regulation may involve inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T-cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active non-antigen-specific process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after the end of exposure to the tolerizing agent. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions, such as, for example, B7 costimulation), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (e.g., B7-1, B7-3) or blocking antibody), prior to transplantation, can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve

sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as described in Lenschow *et al., Science* 257:789-792, 1992 and Turka *et al., Proc. Natl. Acad. Sci USA*, 89:11102-11105, 1992. In addition, murine models of GVHD (see Paul ed., *Fundamental Immunology*, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor/ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which potentially involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/pr/pr mice or NZB hybrid mice, murine autoimmuno collagen arthritis, diabetes mellitus in OD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., *supra*, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may involve either enhancing an existing immune response or eliciting an initial immune response as shown by the following examples. For instance, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases

of viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory form of B lymphocyte antigens systemically.

Alternatively, antiviral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide encoded by extended cDNAs derived from the 5' ESTs of the present invention or together with a stimulatory form of a soluble peptide encoded by extended cDNAs derived from the 5' ESTs of the present invention and reintroducing the *in vitro* primed T cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to T cells *in vivo*, thereby activating the T cells.

In another application, upregulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (e.g., sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide encoded by extended cDNAs derived from the 5' ESTs of the present invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection *in vivo*.

The presence of the peptide encoded by extended cDNAs derived from the 5' ESTs of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules can be transfected with nucleic acids encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I α chain and β_2 microglobulin or an MHC class II α chain and an MHC class II β chain to thereby express MHC class I or MHC class II proteins on the cell surface, respectively. Expression of the appropriate MHC class I or class II

molecules in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject. Alternatively, as described in more detail below, genes encoding these immune system regulator proteins or nucleic acids regulating the expression of such proteins may be introduced into appropriate host cells to increase or decrease the expression of the proteins as desired.

EXAMPLE 34

Assaying the Proteins Expressed from Extended cDNAs or Portions Thereof for Hematopoiesis Regulating Activity

The proteins encoded by the extended cDNAs or portions thereof may also be evaluated for their hematopoiesis regulating activity. For example, the effect of the proteins on embryonic stem cell differentiation may be evaluated. Numerous assays for such activity are familiar to those skilled in the art, including the assays disclosed in the following references, which are incorporated herein by reference: Johansson *et al. Cell. Biol.* **15**:141-151, 1995; Keller *et al., Mol. Cell. Biol.* **13**:473-486, 1993; McClanahan *et al., Blood* **81**:2903-2915, 1993.

The proteins encoded by the extended cDNAs or portions thereof may also be evaluated for their influence on the lifetime of stem cells and stem cell differentiation. Numerous assays for such activity are familiar to those skilled in the art, including the assays disclosed in the following references, which are incorporated herein by reference: Freshney, Methylcellulose Colony Forming Assays, in *Culture of Hematopoietic Cells.*, Freshney, *et al.*. Eds. pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama *et al., Proc. Natl. Acad. Sci. USA* **89**:5907-5911, 1992; McNiece and Briddell, in *Culture of Hematopoietic Cells, supra*; Neben *et al., Exp. Hematol.* **22**:353-359, 1994; Ploemacher and Cobblestone In

*Culture of Hematopoietic Cells, supra*1-21, Spooncer *et al*, in *Culture of Hematopoietic Cells, supra*163-179 and Sutherland in *Culture of Hematopoietic Cells, supra*. 139-162.

Those proteins which exhibit hematopoiesis regulatory activity may then be formulated as pharmaceuticals and used to treat clinical conditions in which regulation of hematopoeisis is beneficial, such as in the treatment of myeloid or lymphoid cell deficiencies. Involvement in regulating hematopoiesis is indicated even by marginal biological activity in support of colony forming cells or of factor-dependent cell lines. For example, proteins supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, indicates utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells. Proteins supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (*i.e.*, traditional CSF activity) may be useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression. Proteins supporting the growth and proliferation of megakaryocytes and consequently of platelets allows prevention or treatment of various platelet disorders such as thrombocytopenia, and generally may be used in place of or complementary to platelet transfusions. Proteins supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells may therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantion, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in vivo* or *ex vivo* (*i.e.*, in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy. Alternatively, as described in more detail below, genes encoding hematopoiesis regulating activity proteins or nucleic acids regulating the expression of such proteins may be introduced into appropriate host cells to increase or decrease the expression of the proteins as desired.

EXAMPLE 35

Assaying the Proteins Expressed from Extended cDNAs or Portions Thereof for Regulation of Tissue Growth

The proteins encoded by the extended cDNAs or portions thereof may also be evaluated for their effect on tissue growth. Numerous assays for such activity are familiar to those skilled in the art, including the assays disclosed in International Patent Publication No. WO95/16035, International Patent Publication No. WO95/05846 and International Patent Publication No. WO91/07491, which are incorporated herein by reference.

Assays for wound healing activity include, without limitation, those described in: Winter, *Epidermal Wound Healing*, pps. 71-112, Maibach and Rovee, eds., Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, *J. Invest. Dermatol.* 71:382-84, 1978, which are incorporated herein by reference.

Those proteins which are involved in the regulation of tissue growth may then be formulated as pharmaceuticals and used to treat clinical conditions in which regulation of tissue growth is beneficial. For example, a protein encoded by extended cDNAs derived from the 5' ESTs of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.

A protein encoded by extended cDNAs derived from the 5' ESTs of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. *De novo* bone synthesis induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of bone-forming cell progenitors. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or

by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein encoded by extended cDNAs derived from the 5' ESTs of the present invention is tendon/ligament formation. A protein encoded by extended cDNAs derived from the 5' ESTs of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. *De novo* tendon/ligament-like tissue formation induced by a composition encoded by extended cDNAs derived from the 5' ESTs of the present invention contributes to the repair of tendon or ligaments defects of congenital, traumatic or other origin and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions encoded by extended cDNAs derived from the 5' ESTs of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein encoded by extended cDNAs derived from the 5' ESTs of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.*, for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and

Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein encoded by extended cDNAs derived from the 5' ESTs of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium) muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to generate. A protein of the invention may also exhibit angiogenic activity.

A protein encoded by extended cDNAs derived from the 5' ESTs of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokinc damage.

A protein encoded by extended cDNAs derived from the 5' ESTs of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

Alternatively, as described in more detail below, genes encoding tissue growth regulating activity proteins or nucleic acids regulating the expression of such proteins may be introduced into appropriate host cells to increase or decrease the expression of the proteins as desired.

EXAMPLE 36

Assaying the Proteins Expressed from Extended cDNAs or Portions Thereof for Regulation of Reproductive Hormones

The proteins encoded by the extended cDNAs or portions thereof may also be evaluated for their ability to regulate reproductive hormones, such as follicle stimulating hormone. Numerous assays for such activity are familiar to those skilled in the art, including the assays disclosed in the following references, which are incorporated herein by reference: Vale *et al.*, *Endocrinol.* **91**:562-572, 1972; Ling *et al.*, *Nature* **321**:779-782, 1986; Vale *et al.*, *Nature* **321**:776-779, 1986; Mason *et al.*., *Nature* **318**:659-663, 1985; Forage *et al.*, *Proc. Natl. Acad. Sci. USA* **83**:3091-3095, 1986, Chapter 6.12 in *Current Protocols in Immunology*, Coligan *et al.* Eds. Greene Publishing Associates and Wiley-Intersciece ; Taub *et al.*, *J. Clin. Invest.* **95**:1370-1376, 1995; Lind *et al.*, *APMIS* **103**:140-146, 1995; Muller *et al.*, *Eur. J. Immunol.* **25**:1744-1748; Gruber *et al.*, *J. Immunol.* **152**:5860-5867, 1994; Johnston *et al.*, *J Immunol.* **153**:1762-1768, 1994.

Those proteins which exhibit activity as reproductive hormones or regulators of cell movement may then be formulated as pharmaceuticals and used to treat clinical conditions in which regulation of reproductive hormones are beneficial. For example, a protein encoded by extended cDNAs derived from the 5' ESTs of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins are characterized by their ability to stimulate the release of FSH. Thus, a protein encoded by extended cDNAs derived from the 5' ESTs of the present invention, alone or in heterodimers with a member of the inhibin α family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin-B group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, United States Patent 4,798,885, the disclosure of which is incorporated herein by reference. A protein of the invention may also be useful for advancement of the onset of

fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

Alternatively, as described in more detail below, genes encoding reproductive hormone regulating activity proteins or nucleic acids regulating the expression of such proteins may be introduced into appropriate host cells to increase or decrease the expression of the proteins as desired.

EXAMPLE 37

Assaying the Proteins Expressed from Extended cDNAs or Portions Thereof for Chemotactic/Chemokinetic Activity

The proteins encoded by the extended cDNAs or portions thereof may also be evaluated for chemotactic/chemokinetic activity. For example, a protein encoded by extended cDNAs derived from the 5' ESTs of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of

cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: *Current Protocols in Immunology*, Ed by Coligan, Kruisbeek, Margulies, Shevach and Strober, Pub. Greene Publishing Associates and Wiley-Interscience, Chapter **6.12**: 6.12.1-6.12.28; Taub *et al., J. Clin. Invest.* **95**:1370-1376, 1995; Lind *et al., APMIS* **103**:140-146, 1995; Mueller *et al., Eur. J. Immunol.* **25**:1744-1748; Gruber *et al., J. Immunol.* **152**:5860-5867, 1994; Johnston *et al. J. Immunol.*, **153**:1762-1768, 1994.

EXAMPLE 38

Assaying the Proteins Expressed from Extended cDNAs or Portions Thereof for Regulation of Blood Clotting

The proteins encoded by the extended cDNAs or portions thereof may also be evaluated for their effects on blood clotting. Numerous assays for such activity are familiar to those skilled in the art, including the assays disclosed in the following references, which are incorporated herein by reference: Linet *et al., J. Clin. Pharmacol.* **26**:131-140, 1986; Burdick *et al., Thrombosis Res.* **45**:413-419, 1987; Humphrey *et al., Fibrinolysis* **5**:71-79, 1991; Schaub, *Prostaglandins* **35**:467-474, 1988.

Those proteins which are involved in the regulation of blood clotting may then be formulated as pharmaceuticals and used to treat clinical conditions in which regulation of blood clotting is beneficial. For example, a protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulations disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as infarction of cardiac and central nervous system vessels (e.g., stroke)). Alternatively, as described in more detail below, genes encoding blood clotting activity proteins or nucleic acids regulating the expression of such proteins may be introduced into appropriate host cells to increase or decrease the expression of the proteins as desired.

EXAMPLE 39

Assaying the Proteins Expressed from Extended cDNAs or Portions Thereof for Involvement in Receptor/Ligand Interactions

The proteins encoded by the extended cDNAs or a portion thereof may also be evaluated for their involvement in receptor/ligand interactions. Numerous assays for such involvement are familiar to those skilled in the art, including the assays disclosed in the following references, which are incorporated herein by reference: Chapter 7. 7.28.1-7.28.22 in *Current Protocols in Immunology*, Coligan *et al.* Eds. Greene Publishing Associates and Wiley-Interscience; Takai *et al., Proc. Natl. Acad. Sci. USA* **84**:6864-6868, 1987; Bierer *et al., J. Exp. Med.* **168**:1145-1156, 1988; Rosenstein *et al., J. Exp. Med.* **169**:149-160, 1989; Stoltenborg *et al., J. Immunol. Methods* 175:59-68, 1994; Stitt *et al., Cell* 80:661-670, 1995; Gyuris *et al., Cell* 75:791-803, 1993.

For example, the proteins encoded by extended cDNAs derived from the 5' ESTs of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein encoded by extended cDNAs derived from the 5' ESTs of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions. Alternatively, as described in more detail below, genes encoding proteins involved in receptor/ligand interactions or nucleic acids regulating the expression of such proteins may be introduced into appropriate host cells to increase or decrease the expression of the proteins as desired.

EXAMPLE 40

Assaying the Proteins Expressed from Extended cDNAs or Portions Thereof for Anti-Inflammatory Activity

The proteins encoded by the extended cDNAs or a portion thereof may also be evaluated for anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat inflammatory conditions including chronic or acute conditions, including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome), ischemia-reperfusioninury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine- or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material. Alternatively, as described in more detail below, genes encoding anti-inflammatory activity proteins or nucleic acids regulating the expression of such proteins may be introduced into appropriate host cells to increase or decrease the expression of the proteins as desired.

EXAMPLE 41

Assaying the Proteins Expressed from Extended cDNAs or Portions Thereof for Tumor Inhibition Activity

The proteins encoded by the extended cDNAs or a portion thereof may also be evaluated for tumor inhibition activity. In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other

factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth. Alternatively, as described in more detail below, genes tumor inhibition activity proteins or nucleic acids regulating the expression of such proteins may be introduced into appropriate host cells to increase or decrease the expression of the proteins as desired.

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein. Alternatively, as described in more detail below, genes encoding proteins involved in any of the above mentioned activities or nucleic acids regulating the expression of such proteins may be introduced into appropriate host cells to increase or decrease the expression of the proteins as desired.

EXAMPLE 42

Identification of Proteins which Interact with Polypeptides Encoded by Extended cDNAs

Proteins which interact with the polypeptides encoded by cDNAs derived from the 5′ ESTs or fragments thereof, such as receptor proteins, may be identified using two hybrid systems such as the Matchmaker Two Hybrid System 2 (Catalog No. K1604-1, Clontech). As described in the manual accompanying the kit which is incorporated herein by reference, the the cDNAs derived from 5′ ESTs, or fragments thereof, are inserted into an expression vector such that they are in frame with DNA encoding the DNA binding domain of the yeast transcriptional activator GAL4. cDNAs in a cDNA library which encode proteins which might interact with the polypeptides encoded by the extended cDNAs or portions thereof are inserted into a second expression vector such that they are in frame with DNA encoding the activation domain of GAL4. The two expression plasmids are transformed into yeast and the yeast are plated on selection medium which selects for expression of selectable markers on each of the expression vectors as well as GAL4 dependent expression of the HIS3 gene. Transformants capable of growing on medium lacking histidine are screened for GAL4 dependent lacZ expression. Those cells which are positive in both the histidine selection and the lacZ assay contain plasmids encoding proteins which interact with the polypeptide encoded by the extended cDNAs or portions thereof.

Alternatively, the system described in Lustig *et al.*, *Methods in Enzymology* **283**: 83-99, 1997, and in U.S. Patent No. 5,654,150, the disclosure of which is incorporated herein by reference, may be used for identifying molecules which interact with the polypeptides encoded by extended cDNAs. In such systems, *in vitro* transcription reactions are performed on a pool of vectors containing extended cDNA inserts cloned downstream of a promoter which drives *in vitro* transcription. The resulting pools of mRNAs are introduced into *Xenopus laevis* oocytes. The oocytes are then assayed for a desired activity.

Alternatively, the pooled *in vitro* transcription products produced as described above may be translated *in vitro*. The pooled *in vitro* translation products can be assayed for a desired activity or for interaction with a known polypeptide.

Proteins or other molecules interacting with polypeptides encoded by extended cDNAs can be found by a variety of additional techniques. In one method, affinity

columns containing the polypeptide encoded by the extended cDNA or a portion thereof can be constructed. In some versions, of this method the affinity column contains chimeric proteins in which the protein encoded by the extended cDNA or a portion thereof is fused to glutathione S-transferase. A mixture of cellular proteins or pool of expressed proteins as described above and is applied to the affinity column. Proteins interacting with the polypeptide attached to the column can then be isolated and analyzed on 2-D electrophoresis gel as described in Ramunsen *et al.*, *Electrophoresis* **18**:588-598, 1997, the disclosure of which is incorporated herein by reference. Alternatively, the proteins retained on the affinity column can be purified by electrophoresis based methods and sequenced. The same method can be used to isolate antibodies, to screen phage display products, or to screen phage display human antibodies.

Proteins interacting with polypeptides encoded by extended cDNAs or portions thereof can also be screened by using an Optical Biosensor as described in Edwards and Leatherbarrow, *Analytical Biochemistry* **246**:1-6, 1997, the disclosure of which is incorporated herein by reference. The main advantage of the method is that it allows the determination of the association rate between the protein and other interacting molecules. Thus, it is possible to specifically select interacting molecules with a high or low association rate. Typically a target molecule is linked to the sensor surface (through a carboxymethl dextran matrix) and a sample of test molecules is placed in contact with the target molecules. The binding of a test molecule to the target molecule causes a change in the refractive index and/ or thickness. This change is detected by the Biosensor provided it occurs in the evanescent field (which extend a few hundred nanometers from the sensor surface). In these screening assays, the target molecule can be one of the polypeptides encoded by extended cDNAs or a portion thereof and the test sample can be a collection of proteins extracted from tissues or cells, a pool of expressed proteins, combinatorial peptide and/ or chemical libraries, or phage displayed peptides. The tissues or cells from which the test proteins are extracted can originate from any species.

In other methods, a target protein is immobilized and the test population is a collection of unique polypeptides encoded by the extended cDNAs or portions thereof.

To study the interaction of the proteins encoded by the extended cDNAs or portions thereof with drugs, the microdialysis coupled to HPLC method described by Wang *et al.*, *Chromatographia* 44:205-208, 1997 or the affinity capillary electrophoresis method described by Busch *et al.*, *J. Chromatogr.* 777:311-328, 1997, the disclosures of which are incorporated herein by reference can be used.

It will be appreciated by those skilled in the art that the proteins expressed from the extended cDNAs or portions may be assayed for numerous activities in addition to those specifically enumerated above. For example, the expressed proteins may be evaluated for applications involving control and regulation of inflammation, tumor proliferation or metastasis, infection, or other clinical conditions. In addition, the proteins expressed from the extended cDNAs or portions thereof may be useful as nutritional agents or cosmetic agents.

The proteins expressed from the cDNAs or portions thereof may be used to generate antibodies capable of specifically binding to the expressed protein or fragments thereof as described in Example 40 below. The antibodies may capable of binding a full length protein encoded by a cDNA derived from a 5′ EST, a mature protein (*i.e.* the protein generated by cleavage of the signal peptide) encoded by a cDNA derived from a 5′ EST, or a signal peptide encoded by a cDNA derived from a 5′ EST. Alternatively, the antibodies may be capable of binding fragments of at least 10 amino acids of the proteins encoded by the above cDNAs. In some embodiments, the antibodies may be capable of binding fragments of at least 15 amino acids of the proteins encoded by the above cDNAs. In other embodiments, the antibodies may be capable of binding fragments of at least 25 amino acids of the proteins expressed from the extended cDNAs which comprise at least 25 amino acids of the proteins encoded by the above cDNAs. In further embodiments, the antibodies may be capable of binding fragments of at least 40 amino acids of the proteins encoded by the above cDNAs.

EXAMPLE 43

Production of an Antibody to a Human Protein

Substantially pure protein or polypeptide is isolated from the transfected or transformed cells as described in Example 30. The concentration of protein in the final preparation is adjusted, for example, by concentration on an Amicon filter device, to the

level of a few µg/ml. Monoclonal or polyclonal antibody to the protein can then be prepared as follows:

1. Monoclonal Antibody Production by Hybridoma Fusion

Monoclonal antibody to epitopes of any of the peptides identified and isolated as described can be prepared from murine hybridomas according to the classical method of Kohler, and Milstein, *Nature* **256**:495, 1975 or derivative methods thereof. Briefly, a mouse is repetitively inoculated with a few micrograms of the selected protein or peptides derived therefrom over a period of a few weeks. The mouse is then sacrificed, and the antibody producing cells of the spleen isolated. The spleen cells are fused by means of polyethylene glycol with mouse myeloma cells, and the excess unfused cells destroyed by growth of the system on selective media comprising aminopterin (HAT media). The successfully fused cells are diluted and aliquots of the dilution placed in wells of a microtiter plate where growth of the culture is continued. Antibody-producing clones are identified by detection of antibody in the supernatant fluid of the wells by immunoassay procedures, such as ELISA, as originally described by Engvall, *Meth. Enzymol.* **70**:419, 1980, the disclosure of which is incorporated herein by reference and derivative methods thereof. Selected positive clones can be expanded and their monoclonal antibody product harvested for use. Detailed procedures for monoclonal antibody production are described in Davis *et al.* in *Basic Methods in Molecular Biology* Elsevier, New York. Section 21-2, the disclosure of which is incorporated herein by reference.

2. Polyclonal Antibody Production by Immunization

Polyclonal antiserum containing antibodies to heterogenous epitopes of a single protein can be prepared by immunizing suitable animals with the expressed protein or peptides derived therefrom, which can be unmodified or modified to enhance immunogenicity. Effective polyclonal antibody production is affected by many factors related both to the antigen and the host species. For example, small molecules tend to be less immunogenic than others and may require the use of carriers and adjuvant. Also, host animals response vary depending on site of inoculations and doses, with both inadequate or

excessive doses of antigen resulting in low titer antisera. Small doses (ng level) of antigen administered at multiple intradermal sites appears to be most reliable. An effective immunization protocol for rabbits can be found in Vaitukaitis. *et al, J. Clin. Endocrinol. Metab.* **33**:988-991 (1971) , the disclosure of which is incorporated herein by reference..

Booster injections can be given at regular intervals, and antiserum harvested when antibody titer thereof, as determined semi-quantitatively, for example, by double immunodiffusion in agar against known concentrations of the antigen, begins to fall. See, for example, Ouchterlony, *et al.*, Chap. 19 in: *Handbook of Experimental Immunology* D. Wier (ed) Blackwell (1973) , the disclosure of which is incorporated herein by reference. Plateau concentration of antibody is usually in the range of 0.1 to 0.2 mg/ml of serum (about 12 μM). Affinity of the antisera for the antigen is determined by preparing competitive binding curves, as described, for example, by Fisher, D., Chap. 42 in: *Manual of Clinical Immunology*, 2d Ed. (Rose and Friedman, Eds.) Amer. Soc. For Microbiol., Washington, D.C. (1980) , the disclosure of which is incorporated herein by reference.

Antibody preparations prepared according to either protocol are useful in quantitative immunoassays which determine concentrations of antigen-bearing substances in biological samples; they are also used semi-quantitatively or qualitatively to identify the presence of antigen in a biological sample. The antibodies may also be used in therapeutic compositions for killing cells expressing the protein or reducing the levels of the protein in the body.

V. Use of 5' ESTs or Sequences Obtainable Therefrom or Portions Thereof as Reagents

The 5′ ESTs of the present invention (or cDNAs or genomic DNAs obtainable therefrom) may be used as reagents in isolation procedures, diagnostic assays, and forensic procedures. For example, sequences from the 5′ ESTs (or cDNAs or genomic DNAs obtainable therefrom) may be detectably labeled and used as probes to isolate other sequences capable of hybridizing to them. In addition, sequences from 5′ ESTs (or cDNAs or genomic DNAs obtainable therefrom) may be used to design PCR primers to be used in isolation, diagnostic, or forensic procedures.

1. Use of 5' ESTs or Sequences Obtainable Therefrom or Portions Thereof in Isolation, Diagnostic and Forensic Procedures

EXAMPLE 44

Preparation of PCR Primers and Amplification of DNA

The 5' EST sequences (or cDNAs or genomic DNAs obtainable therefrom) may be used to prepare PCR primers for a variety of applications, including isolation procedures for cloning nucleic acids capable of hybridizing to such sequences, diagnostic techniques and forensic techniques. The PCR primers are at least 10 bases, and preferably at least 12, 15, or 17 bases in length. More preferably, the PCR primers are at least 20-30 bases in length. In some embodiments, the PCR primers may be more than 30 bases in length. It is preferred that the primer pairs have approximately the same G/C ratio, so that melting temperatures are approximately the same. A variety of PCR techniques are familiar to those skilled in the art. For a review of PCR technology, see Molecular Cloning to Genetic Engineering, White Ed. *in Methods in Molecular Biology* 67: Humana Press, Totowa 1997, the disclosure of which is incorporated herein by reference. In each of these PCR procedures, PCR primers on either side of the nucleic acid sequences to be amplified are added to a suitably prepared nucleic acid sample along with dNTPs and a thermostable polymerase such as Taq polymerase, Pfu polymerase, or Vent polymerase. The nucleic acid in the sample is denatured and the PCR primers are specifically hybridized to complementary nucleic acid sequences in the sample. The hybridized primers are extended. Thereafter, another cycle of denaturation, hybridization, and extension is initiated. The cycles are repeated multiple times to produce an amplified fragment containing the nucleic acid sequence between the primer sites.

EXAMPLE 45

Use of 5'ESTs as Probes

Probes derived from 5' ESTs (or cDNAs or genomic DNAs obtainable therefrom), including full length cDNAs or genomic sequences, may be labeled with detectable labels familiar to those skilled in the art, including radioisotopes and non-radioactive labels, to provide a detectable probe. The detectable probe may be single stranded or double stranded and may be made using techniques known in the art, including *in vitro* transcription, nick translation, or kinase reactions. A nucleic acid sample containing a sequence capable of

hybridizing to the labeled probe is contacted with the labeled probe. If the nucleic acid in the sample is double stranded, it may be denatured prior to contacting the probe. In some applications, the nucleic acid sample may be immobilized on a surface such as a nitrocellulose or nylon membrane. The nucleic acid sample may comprise nucleic acids obtained from a variety of sources, including genomic DNA, cDNA libraries, RNA, or tissue samples.

Procedures used to detect the presence of nucleic acids capable of hybridizing to the detectable probe include well known techniques such as Southern blotting, Northern blotting, dot blotting, colony hybridization, and plaque hybridization. In some applications, the nucleic acid capable of hybridizing to the labeled probe may be cloned into vectors such as expression vectors, sequencing vectors, or *in vitro* transcription vectors to facilitate the characterization and expression of the hybridizing nucleic acids in the sample. For example, such techniques may be used to isolate and clone sequences in a genomic library or cDNA library which are capable of hybridizing to the detectable probe as described in Example 30 above.

PCR primers made as described in Example 44 above may be used in forensic analyses, such as the DNA fingerprinting techniques described in Examples 46-50 below. Such analyses may utilize detectable probes or primers based on the sequences of the the 5′ ESTs or of cDNAs or genomic DNAs isolated using the 5′ ESTs.

EXAMPLE 46

Forensic Matching by DNA Sequencing

In one exemplary method, DNA samples are isolated from forensic specimens of, for example, hair, semen, blood or skin cells by conventional methods. A panel of PCR primers based on a number of the 5′ ESTs of Example 25, or cDNAs or genomic DNAs isolated therefrom as described above, is then utilized in accordance with Example 44 to amplify DNA of approximately 100-200 bases in length from the forensic specimen. Corresponding sequences are obtained from a test subject. Each of these identification DNAs is then sequenced using standard techniques, and a simple database comparison determines the differences, if any, between the sequences from the subject and those from the sample. Statistically significant differences between the suspect's DNA sequences and those from the sample conclusively prove a lack of identity. This lack of identity can be proven, for example, with only one sequence. Identity, on the other hand, should be demonstrated with a large

number of sequences, all matching. Preferably, a minimum of 50 statistically identical sequences of 100 bases in length are used to prove identity between the suspect and the sample.

EXAMPLE 47

Positive Identification by DNA Sequencing

The technique outlined in the previous example may also be used on a larger scale to provide a unique fingerprint-type identification of any individual. In this technique, primers are prepared from a large number of 5'EST sequences from Example 25, or cDNA or genomic DNA sequences obtainable therefrom. Preferably, 20 to 50 different primers are used. These primers are used to obtain a corresponding number of PCR-generated DNA segments from the individual in question in accordance with Example 44. Each of these DNA segments is sequenced, using the methods set forth in Example 46. The database of sequences generated through this procedure uniquely identifies the individual from whom the sequences were obtained. The same panel of primers may then be used at any later time to absolutely correlate tissue or other biological specimen with that individual.

EXAMPLE 48

Southern Blot Forensic Identification

The procedure of Example 47 is repeated to obtain a panel of at least 10 amplified sequences from an individual and a specimen. Preferably, the panel contains at least 50 amplified sequences. More preferably, the panel contains 100 amplified sequences. In some embodiments, the panel contains 200 amplified sequences. This PCR-generated DNA is then digested with one or a combination of, preferably, four base specific restriction enzymes. Such enzymes are commercially available and known to those of skill in the art. After digestion, the resultant gene fragments are size separated in multiple duplicate wells on an agarose gel and transferred to nitrocellulose using Southern blotting techniques well known to those with skill in the art. For a review of Southern blotting see Davis *et al.* (Basic Methods in Molecular Biology, 1986, Elsevier Press. pp 62-65) , the disclosure of which is incorporated herein by reference.

A panel of probes based on the sequences of 5′ ESTs (or cDNAs or genomic DNAs obtainable therefrom), or fragments thereof of at least 10 bases, are radioactively or colorimetrically labeled using methods known in the art, such as nick translation or end labeling, and hybridized to the Southern blot using techniques known in the art (Davis *et al.*, supra). Preferably, the probe comprises at least 12, 15, or 17 consecutive nucleotides from the 5′ EST (or cDNAs or genomic DNAs obtainable therefrom). More preferably, the probe comprises at least 20-30 consecutive nucleotides from the 5′ EST (or cDNAs or genomic DNAs obtainable therefrom). In some embodiments, the probe comprises more than 30 nucleotides from the 5′ EST (or cDNAs or genomic DNAs obtainable therefrom).

Preferably, at least 5 to 10 of these labeled probes are used, and more preferably at least about 20 or 30 are used to provide a unique pattern. The resultant bands appearing from the hybridization of a large sample of 5′ EST (or cDNAs or genomic DNAs obtainable therefrom) will be a unique identifier. Since the restriction enzyme cleavage will be different for every individual, the band pattern on the Southern blot will also be unique. Increasing the number of 5′ EST (or cDNAs or genomic DNAs obtainable therefrom) probes will provide a statistically higher level of confidence in the identification since there will be an increased number of sets of bands used for identification.

EXAMPLE 49

Dot Blot Identification Procedure

Another technique for identifying individuals using the 5′ EST sequences disclosed herein utilizes a dot blot hybridization technique.

Genomic DNA is isolated from nuclei of subject to be identified. Oligonucleotide probes of approximately 30 bp in length are synthesized that correspond to at least 10, preferably 50 sequences from the 5′ ESTs or cDNAs or genomic DNAs obtainable therefrom. The probes are used to hybridize to the genomic DNA through conditions known to those in the art. The oligonucleotides are end labeled with P^{32} using polynucleotide kinase (Pharmacia). Dot Blots are created by spotting the genomic DNA onto nitrocellulose or the like using a vacuum dot blot manifold (BioRad, Richmond California). The nitrocellulose filter containing the genomic sequences is baked or UV linked to the filter, prehybridized and hybridized with labeled probe using techniques known in the art (Davis et *al.*, *supra*). The ^{32}P

labeled DNA fragments are sequentially hybridized with successively stringent conditions to detect minimal differences between the 30 bp sequence and the DNA. Tetramethylammonium chloride is useful for identifying clones containing small numbers of nucleotide mismatches (Wood *et al.*, *Proc. Natl. Acad. Sci. USA* **82(6)**:1585-1588, 1985) which is hereby incorporated by reference. A unique pattern of dots distinguishes one individual from another individual.

5′ EST sequences (or cDNAs or genomic DNAs obtainable therefrom) or oligonucleotides containing at least 10 consecutive bases from these sequences can be used as probes in the following alternative fingerprinting technique. Preferably, the probe comprises at least 12, 15, or 17 consecutive nucleotides from the 5′ EST sequences (or cDNAs or genomic DNAs obtainable therefrom). More preferably, the probe comprises at least 20-30 consecutive nucleotides from the 5′ EST sequences (or cDNAs or genomic DNAs obtainable therefrom). In some embodiments, the probe comprises more than 30 nucleotides from the 5′ EST sequences (or cDNAs or genomic DNAs obtainable therefrom).

Preferably, a plurality of probes having sequences from different genes are used in the alternative fingerprinting technique. Example 50 below provides a representative alternative fingerprinting procedure in which the probes are derived from 5'EST.

EXAMPLE 50

Alternative "Fingerprint" Identification Technique

20-mer oligonucleotides are prepared from a large number, e.g. 50, 100, or 200, of 5'EST using commercially available oligonucleotide services such as Genset, Paris, France. Cell samples from the test subject are processed for DNA using techniques well known to those with skill in the art. The nucleic acid is digested with restriction enzymes such as EcoRI and XbaI. Following digestion, samples are applied to wells for electrophoresis. The procedure, as known in the art, may be modified to accommodate polyacrylamide electrophoresis, however in this example, samples containing 5 ug of DNA are loaded into wells and separated on 0.8% agarose gels. The gels are transferred onto nitrocellulose using standard Southern blotting techniques.

10 ng of each of the oligonucleotides are pooled and end-labeled with ^{32}P. The nitrocellulose is prehybridized with blocking solution and hybridized with the labeled probes.

Following hybridization and washing, the nitrocellulose filter is exposed to X-Omat AR X-ray film. The resulting hybridization pattern will be unique for each individual.

It is additionally contemplated within this example that the number of probe sequences used can be varied for additional accuracy or clarity.

The proteins encoded by the extended cDNAs may also be used to generate antibodies as explained in Examples 30 and 43 in order to identify the tissue type or cell species from which a sample is derived as described in example 51.

EXAMPLE 51

Identification of Tissue Types or Cell Species by Means of Labeled Tissue Specific Antibodies

Identification of specific tissues is accomplished by the visualization of tissue specific antigens by means of antibody preparations according to Examples 30 and 43 which are conjugated, directly or indirectly to a detectable marker. Selected labeled antibody species bind to their specific antigen binding partner in tissue sections, cell suspensions, or in extracts of soluble proteins from a tissue sample to provide a pattern for qualitative or semi-qualitative interpretation.

Antisera for these procedures must have a potency exceeding that of the native preparation, and for that reason, antibodies are concentrated to a mg/ml level by isolation of the gamma globulin fraction, for example, by ion-exchange chromatography or by ammonium sulfate fractionation. Also, to provide the most specific antisera, unwanted antibodies, for example to common proteins, must be removed from the gamma globulin fraction, for example by means of insoluble immunoabsorbents, before the antibodies are labeled with the marker. Either monoclonal or heterologous antisera is suitable for either procedure.

A. Immunohistochemical techniques

Purified, high-titer antibodies, prepared as described above, are conjugated to a detectable marker, as described, for example, by Fudenberg, Chap. 26 in: *Basic and Clinical Immunology*, 3rd Ed. Lange, Los Altos, California, 1980, or Rose,. *et al.*, Chap. **12** in:

Methods in Immunodiagnosis, 2d Ed. John Wiley and Sons, New York (1980), the disclosures of which are incorporated herein by reference.

A fluorescent marker, either fluorescein or rhodamine, is preferred, but antibodies can also be labeled with an enzyme that supports a color producing reaction with a substrate, such as horseradish peroxidase. Markers can be added to tissue-bound antibody in a second step, as described below. Alternatively, the specific antitissue antibodies can be labeled with ferritin or other electron dense particles, and localization of the ferritin coupled antigen-antibody complexes achieved by means of an electron microscope. In yet another approach, the antibodies are radiolabeled, with, for example ^{125}I, and detected by overlaying the antibody treated preparation with photographic emulsion.

Preparations to carry out the procedures can comprise monoclonal or polyclonal antibodies to a single protein or peptide identified as specific to a tissue type, for example, brain tissue, or antibody preparations to several antigenically distinct tissue specific antigens can be used in panels, independently or in mixtures, as required.

Tissue sections and cell suspensions are prepared for immunohistochemical examination according to common histological techniques. Multiple cryostat sections (about 4 μm, unfixed) of the unknown tissue and known control, are mounted and each slide covered with different dilutions of the antibody preparation. Sections of known and unknown tissues should also be treated with preparations to provide a positive control, a negative control, for example, pre-immune sera, and a control for non-specific staining, for example, buffer.

Treated sections are incubated in a humid chamber for 30 min at room temperature, rinsed, then washed in buffer for 30-45 min. Excess fluid is blotted away, and the marker developed.

If the tissue specific antibody was not labeled in the first incubation, it can be labeled at this time in a second antibody-antibody reaction, for example, by adding fluorescein- or enzyme-conjugated antibody against the immunoglobulin class of the antiserum-producing species, for example, fluorescein labeled antibody to mouse IgG. Such labeled sera are commercially available.

The antigen found in the tissues by the above procedure can be quantified by measuring the intensity of color or fluorescence on the tissue section, and calibrating that signal using appropriate standards.

B. Identification of tissue specific soluble proteins

The visualization of tissue specific proteins and identification of unknown tissues from that procedure is carried out using the labeled antibody reagents and detection strategy as described for immunohistochemistry; however the sample is prepared according to an electrophoretic technique to distribute the proteins extracted from the tissue in an orderly array on the basis of molecular weight for detection.

A tissue sample is homogenized using a Virtis apparatus; cell suspensions are disrupted by Dounce homogenization or osmotic lysis, using detergents in either case as required to disrupt cell membranes, as is the practice in the art. Insoluble cell components such as nuclei, microsomes, and membrane fragments are removed by ultracentrifugation, and the soluble protein-containing fraction concentrated if necessary and reserved for analysis.

A sample of the soluble protein solution is resolved into individual protein species by conventional SDS polyacrylamide electrophoresis as described, for example, by Davis, *et al.*, Section 19-2 in: *Basic Methods in Molecular Biology*, Leder ed., Elsevier, New York, 1986, the disclosure of which is incorporated herein by reference, using a range of amounts of polyacrylamide in a set of gels to resolve the entire molecular weight range of proteins to be detected in the sample. A size marker is run in parallel for purposes of estimating molecular weights of the constituent proteins. Sample size for analysis is a convenient volume of from 5 to 55 μl, and containing from about 1 to 100 μg protein. An aliquot of each of the resolved proteins is transferred by blotting to a nitrocellulose filter paper, a process that maintains the pattern of resolution. Multiple copies are prepared. The procedure, known as Western Blot Analysis, is well described in Davis, L. *et al., supra* Section 19-3. One set of nitrocellulose blots is stained with Coomassie blue dye to visualize the entire set of proteins for comparison with the antibody bound proteins. The remaining nitrocellulose filters are then incubated with a solution of one or more specific antisera to tissue specific proteins prepared as described in Examples 30 and 43. In this procedure, as in procedure A above, appropriate positive and negative sample and reagent controls are run.

In either procedure A or B, a detectable label can be attached to the primary tissue antigen-primary antibody complex according to various strategies and permutations thereof. In a straightforward approach, the primary specific antibody can be labeled; alternatively, the unlabeled complex can be bound by a labeled secondary anti-IgG antibody. In other approaches, either the primary or secondary antibody is conjugated to a biotin molecule, which can, in a subsequent step, bind an avidin conjugated marker. According to yet another strategy, enzyme labeled or radioactive protein A, which has the property of binding to any IgG, is bound in a final step to either the primary or secondary antibody.

The visualization of tissue specific antigen binding at levels above those seen in control tissues to one or more tissue specific antibodies, prepared from the gene sequences identified from extended cDNA sequences, can identify tissues of unknown origin, for example, forensic samples, or differentiated tumor tissue that has metastasized to foreign bodily sites.

In addition to their applications in forensics and identification, 5′ ESTs (or cDNAs or genomic DNAs obtainable therefrom) may be mapped to their chromosomal locations. Example 52 below describes radiation hybrid (RH) mapping of human chromosomal regions using 5'ESTs. Example 53 below describes a representative procedure for mapping an 5′ EST to its location on a human chromosome. Example 54 below describes mapping of 5′ ESTs on metaphase chromosomes by Fluorescence In Situ Hybridization (FISH). Those skilled in the art will appreciate that the method of Examples 52-54 may also be used to map cDNAs or genomic DNAs obtainable from the 5′ ESTs to their chromosomal locations.

2. Use of 5' ESTs or Sequences Obtainable Therefrom or Portions Thereof in Chromosome Mapping

EXAMPLE 52

Radiation hybrid mapping of 5'ESTs to the human genome

Radiation hybrid (RH) mapping is a somatic cell genetic approach that can be used for high resolution mapping of the human genome. In this approach, cell lines containing one or more human chromosomes are lethally irradiated, breaking each chromosome into fragments whose size depends on the radiation dose. These fragments are rescued by fusion

with cultured rodent cells, yielding subclones containing different portions of the human genome. This technique is described by Benham *et al.*, *Genomics* **4**:509-517, 1989; and Cox *et al.*, *Science* **250**:245-250, 1990, the entire contents of which are hereby incorporated by reference. The random and independent nature of the subclones permits efficient mapping of any human genome marker. Human DNA isolated from a panel of 80-100 cell lines provides a mapping reagent for ordering 5'EST. In this approach, the frequency of breakage between markers is used to measure distance, allowing construction of fine resolution maps as has been done using conventional ESTs (Schuler *et al.*, *Science* **274**:540-546, 1996, hereby incorporated by reference).

RH mapping has been used to generate a high-resolution whole genome radiation hybrid map of human chromosome 17q22-q25.3 across the genes for growth hormone (GH) and thymidine kinase (TK) (Foster *et al.*, *Genomics* **33**:185-192, 1996), the region surrounding the Gorlin syndrome gene (Obermayr *et al.*, *Eur. J. Hum. Genet.* **4**:242-245, 1996), 60 loci covering the entire short arm of chromosome 12 (Raeymaekers *et al.*, *Genomics* **29**:170-178, 1995), the region of human chromosome 22 containing the neurofibromatosis type 2 locus (Frazer *et al.*, *Genomics* **14**:574-584, 1992) and 13 loci on the long arm of chromosome 5 (Warrington *et al.*, *Genomics* **11**:701-708, 1991).

EXAMPLE 53

Mapping of 5'ESTs to HumanChromosomes using PCR techniques

5' ESTs (or cDNAs or genomic DNAs obtainable therefrom) may be assigned to human chromosomes using PCR based methodologies. In such approaches, oligonucleotide primer pairs are designed from the 5' ESTs (or cDNAs or genomic DNAs obtainable therefrom) to minimize the chance of amplifying through an intron. Preferably, the oligonucleotide primers are 18-23 bp in length and are designed for PCR amplification. The creation of PCR primers from known sequences is well known to those with skill in the art. For a review of PCR technology see Erlich in *PCR Technology, Principles and Applications for DNA Amplification*, Freeman and Co., New York, 1992, the disclosure of which is incorporated herein by reference.

The primers are used in polymerase chain reactions (PCR) to amplify templates from total human genomic DNA. PCR conditions are as follows: 60 ng of genomic DNA is used

as a template for PCR with 80 ng of each oligonucleotide primer, 0.6 unit of Taq polymerase, and 1 μCu of a ^{32}P-labeled deoxycytidine triphosphate. The PCR is performed in a microplate thermocycler (Techne) under the following conditions: 30 cycles of 94°C, 1.4 min; 55°C, 2 min; and 72°C, 2 min; with a final extension at 72°C for 10 min. The amplified products are analyzed on a 6% polyacrylamide sequencing gel and visualized by autoradiography. If the length of the resulting PCR product is identical to the distance between the ends of the primer sequences in the extended cDNA from which the primers are derived, then the PCR reaction is repeated with DNA templates from two panels of human-rodent somatic cell hybrids, BIOS PCRable DNA (BIOS Corporation) and NIGMS Human-Rodent Somatic Cell Hybrid Mapping Panel Number 1 (NIGMS, Camden, NJ).

PCR is used to screen a series of somatic cell hybrid cell lines containing defined sets of human chromosomes for the presence of a given 5′ EST (or cDNA or genomic DNA obtainable therefrom). DNA is isolated from the somatic hybrids and used as starting templates for PCR reactions using the primer pairs from the 5′ EST (or cDNA or genomic DNA obtainable therefrom). Only those somatic cell hybrids with chromosomes containing the human gene corresponding to the 5′ EST (or cDNA or genomic DNA obtainable therefrom) will yield an amplified fragment. The 5′ EST (or cDNA or genomic DNA obtainable therefrom) are assigned to a chromosome by analysis of the segregation pattern of PCR products from the somatic hybrid DNA templates. The single human chromosome present in all cell hybrids that give rise to an amplified fragment is the chromosome containing that 5'EST (or cDNA or genomic DNA obtainable therefrom). For a review of techniques and analysis of results from somatic cell gene mapping experiments, see Ledbetter *et al.*, *Genomics* 6:475-481, 1990, the disclosure of which is incorporated herein by reference.

EXAMPLE 54

Mapping of Extended 5′ ESTs to Chromosomes Using Fluorescence *In Situ* Hybridization

Fluorescence in situ hybridization allows the 5'EST (or cDNA or genomic DNA obtainable therefrom) to be mapped to a particular location on a given chromosome. The chromosomes to be used for fluorescence in situ hybridization techniques may be obtained from a variety of sources including cell cultures, tissues, or whole blood.

In a preferred embodiment, chromosomal localization of an 5'EST (or cDNA or genomic DNA obtainable therefrom) is obtained by FISH as described by Cherif *et al.* (*Proc. Natl. Acad. Sci. U.S.A.*, 87:6639-6643, 1990), the disclosure of which is incorporated herein by reference.. Metaphase chromosomes are prepared from phytohemagglutinin (PHA)-stimulated blood cell donors. PHA-stimulated lymphocytes from healthy males are cultured for 72 h in RPMI-1640 medium. For synchronization, methotrexate (10 μM) is added for 17 h, followed by addition of 5-bromodeoxyuridine (5-BrdU, 0.1 mM) for 6 h. Colcemid (1 μg/ml) is added for the last 15 min before harvesting the cells. Cells are collected, washed in RPMI, incubated with a hypotonic solution of KCl (75 mM) at 37°C for 15 min and fixed in three changes of methanol:acetic acid (3:1). The cell suspension is dropped onto a glass slide and air dried. The 5'EST (or cDNA or genomic DNA obtainable therefrom) is labeled with biotin-16 dUTP by nick translation according to the manufacturer's instructions (Bethesda Research Laboratories, Bethesda, MD), purified using a Sephadex G-50 column (Pharmacia, Upsala, Sweden) and precipitated. Just prior to hybridization, the DNA pellet is dissolved in hybridization buffer (50% formamide, 2 X SSC, 10% dextran sulfate, 1 mg/ml sonicated salmon sperm DNA, pH 7) and the probe is denatured at 70°C for 5-10 min.

Slides kept at -20°C are treated for 1 h at 37°C with RNase A (100 μg/ml), rinsed three times in 2 X SSC and dehydrated in an ethanol series. Chromosome preparations are denatured in 70% formamide, 2 X SSC for 2 min at 70°C, then dehydrated at 4°C. The slides are treated with proteinase K (10 μg/100 ml in 20 mM Tris-HCl, 2 mM $CaCl_2$) at 37°C for 8 min and dehydrated. The hybridization mixture containing the probe is placed on the slide, covered with a coverslip, sealed with rubber cement and incubated overnight in a humid chamber at 37°C. After hybridization and post-hybridization washes, the biotinylated probe is detected by avidin-FITC and amplified with additional layers of biotinylated goat anti-avidin and avidin-FITC. For chromosomal localization, fluorescent R-bands are obtained as previously described (Cherif *et al.*, *supra.*). The slides are observed under a LEICA fluorescence microscope (DMRXA). Chromosomes are counterstained with propidium iodide and the fluorescent signal of the probe appears as two symmetrical yellow-green spots on both chromatids of the fluorescent R-band chromosome (red). Thus, a particular 5'EST (or cDNA or genomic DNA obtainable therefrom) may be localized to a particular cytogenetic R-band on a given chromosome.

Once the 5'EST (or cDNA or genomic DNA obtainable therefrom) have been assigned to particular chromosomes using the techniques described in Examples 52-54 above, they may be utilized to construct a high resolution map of the chromosomes on which they are located or to identify the chromosomes in a sample.

EXAMPLE 55

Use of 5'EST to Construct or Expand Chromosome Maps

Chromosome mapping involves assigning a given unique sequence to a particular chromosome as described above. Once the unique sequence has been mapped to a given chromosome, it is ordered relative to other unique sequences located on the same chromosome. One approach to chromosome mapping utilizes a series of yeast artificial chromosomes (YACs) bearing several thousand long inserts derived from the chromosomes of the organism from which the extended cDNAs (or genomic DNAs obtainable therefrom) are obtained. This approach is described in Nagaraja *et al., Genome Research* 7:210-222, 1997, the disclosure of which is incorporated herein by reference. Briefly, in this approach each chromosome is broken into overlapping pieces which are inserted into the YAC vector. The YAC inserts are screened using PCR or other methods to determine whether they include the 5'EST (or cDNA or genomic DNA obtainable therefrom) whose position is to be determined. Once an insert has been found which includes the 5'EST (or cDNA or genomic DNA obtainable therefrom), the insert can be analyzed by PCR or other methods to determine whether the insert also contains other sequences known to be on the chromosome or in the region from which the 5'EST (or cDNA or genomic DNA obtainable therefrom) was derived. This process can be repeated for each insert in the YAC library to determine the location of each of the extended cDNAs (or genomic DNAs obtainable therefrom) relative to one another and to other known chromosomal markers. In this way, a high resolution map of the distribution of numerous unique markers along each of the organisms chromosomes may be obtained.

As described in Example 56 below extended cDNAs (or genomic DNAs obtainable therefrom) may also be used to identify genes associated with a particular phenotype, such as hereditary disease or drug response.

3. Use of 5'ESTs or Sequences Obtained Therefrom or Fragments Thereof in Gene Identification

EXAMPLE 56

Identification of genes associated with hereditary diseases or drug response

This example illustrates an approach useful for the association of 5'ESTs (or cDNA or genomic DNA obtainable therefrom) with particular phenotypic characteristics. In this example, a particular 5'EST (or cDNA or genomic DNA obtainable therefrom) is used as a test probe to associate that 5'EST (or cDNA or genomic DNA obtainable therefrom) with a particular phenotypic characteristic.

5'ESTs (or cDNA or genomic DNA obtainable therefrom) are mapped to a particular location on a human chromosome using techniques such as those described in Examples 52 and 53 or other techniques known in the art. A search of Mendelian Inheritance in Man (McKusick in *Mendelian Inheritance in Man* (available on line through Johns Hopkins University Welch Medical Library) reveals the region of the human chromosome which contains the 5'EST (or cDNA or genomic DNA obtainable therefrom) to be a very gene rich region containing several known genes and several diseases or phenotypes for which genes have not been identified. The gene corresponding to this 5'EST (or cDNA or genomic DNA obtainable therefrom) thus becomes an immediate candidate for each of these genetic diseases.

Cells from patients with these diseases or phenotypes are isolated and expanded in culture. PCR primers from the 5'EST (or cDNA or genomic DNA obtainable therefrom) are used to screen genomic DNA, mRNA or cDNA obtained from the patients. 5'ESTs (or cDNA or genomic DNA obtainable therefrom) that are not amplified in the patients can be positively associated with a particular disease by further analysis. Alternatively, the PCR analysis may yield fragments of different lengths when the samples are derived from an individual having the phenotype associated with the disease than when the sample is derived from a healthy individual, indicating that the gene containing the 5'EST may be responsible for the genetic disease.

VI. Use of 5'EST (or cDNA or Genomic DNA Obtainable Therefrom) to Construct Vectors

The present 5'ESTs (or cDNA or genomic DNA obtainable therefrom) may also be used to construct secretion vectors capable of directing the secretion of the proteins encoded by genes therein. Such secretion vectors may facilitate the purification or enrichment of the proteins encoded by genes inserted therein by reducing the number of background proteins from which the desired protein must be purified or enriched. Exemplary secretion vectors are described in Example 57 below.

1. Construction of Secretion Vectors

EXAMPLE 57

Construction of Secretion Vectors

The secretion vectors include a promoter capable of directing gene expression in the host cell, tissue, or organism of interest. Such promoters include the Rous Sarcoma Virus promoter, the SV40 promoter, the human cytomegalovirus promoter, and other promoters familiar to those skilled in the art.

A signal sequence from a 5' EST (or cDNAs or genomic DNAs obtainable therefrom) is operably linked to the promoter such that the mRNA transcribed from the promoter will direct the translation of the signal peptide. The host cell, tissue, or organism may be any cell, tissue, or organism which recognizes the signal peptide encoded by the signal sequence in the 5' EST (or cDNA or genomic DNA obtainable therefrom). Suitable hosts include mammalian cells, tissues or organisms, avian cells, tissues, or organisms, insect cells, tissues or organisms, or yeast.

In addition, the secretion vector contains cloning sites for inserting genes encoding the proteins which are to be secreted. The cloning sites facilitate the cloning of the insert gene in frame with the signal sequence such that a fusion protein in which the signal peptide is fused to the protein encoded by the inserted gene is expressed from the mRNA transcribed from the promoter. The signal peptide directs the extracellular secretion of the fusion protein.

The secretion vector may be DNA or RNA and may integrate into the chromosome of the host, be stably maintained as an extrachromosomal replicon in the host, be an artificial chromosome, or be transiently present in the host. Many nucleic acid backbones suitable for

use as secretion vectors are known to those skilled in the art, including retroviral vectors, SV40 vectors, Bovine Papilloma Virus vectors, yeast integrating plasmids, yeast episomal plasmids, yeast artificial chromosomes, human artificial chromosomes, P element vectors, baculovirus vectors, or bacterial plasmids capable of being transiently introduced into the host.

The secretion vector may also contain a polyA signal such that the polyA signal is located downstream of the gene inserted into the secretion vector.

After the gene encoding the protein for which secretion is desired is inserted into the secretion vector, the secretion vector is introduced into the host cell, tissue, or organism using calcium phosphate precipitation, DEAE-Dextran, electroporation, liposome-mediated transfection, viral particles or as naked DNA. The protein encoded by the inserted gene is then purified or enriched from the supernatant using conventional techniques such as ammonium sulfate precipitation, immunoprecipitation, immunochromatography, size exclusion chromatography, ion exchange chromatography, and HPLC. Alternatively, the secreted protein may be in a sufficiently enriched or pure state in the supernatant or growth media of the host to permit it to be used for its intended purpose without further enrichment.

The signal sequences may also be inserted into vectors designed for gene therapy. In such vectors, the signal sequence is operably linked to a promoter such that mRNA transcribed from the promoter encodes the signal peptide. A cloning site is located downstream of the signal sequence such that a gene encoding a protein whose secretion is desired may readily be inserted into the vector and fused to the signal sequence. The vector is introduced into an appropriate host cell. The protein expressed from the promoter is secreted extracellularly, thereby producing a therapeutic effect.

The 5′ ESTs may also be used to clone sequences located upstream of the 5′ ESTs which are capable of regulating gene expression, including promoter sequences, enhancer sequences, and other upstream sequences which influence transcription or translation levels. Once identified and cloned, these upstream regulatory sequences may be used in expression vectors designed to direct the expression of an inserted gene in a desired spatial, temporal, developmental, or quantitative fashion. Example 58 describes a method for cloning sequences upstream of the extended cDNAs or 5′ ESTs.

2. Identification of Upstream Sequences With Promoting or Regulatory Activities

EXAMPLE 58

Use of Extended cDNAs or 5′ ESTs to Clone Upstream Sequences from Genomic DNA

Sequences derived from extended cDNAs or 5′ ESTs may be used to isolate the promoters of the corresponding genes using chromosome walking techniques. In one chromosome walking technique, which utilizes the GenomeWalker™ kit available from Clontech, five complete genomic DNA samples are each digested with a different restriction enzyme which has a 6 base recognition site and leaves a blunt end. Following digestion, oligonucleotide adapters are ligated to each end of the resulting genomic DNA fragments.

For each of the five genomic DNA libraries, a first PCR reaction is performed according to the manufacturer's instructions (which are incorporated herein by reference) using an outer adaptor primer provided in the kit and an outer gene specific primer. The gene specific primer should be selected to be specific for the extended cDNA or 5′ EST of interest and should have a melting temperature, length, and location in the extended cDNA or 5'EST which is consistent with its use in PCR reactions. Each first PCR reaction contains 5 ng of genomic DNA, 5 μl of 10X Tth reaction buffer, 0.2 mM of each dNTP, 0.2 μM each of outer adaptor primer and outer gene specific primer, 1.1 mM of $Mg(OAc)_2$, and 1 μl of the Tth polymerase 50X mix in a total volume of 50 μl. The reaction cycle for the first PCR reaction is as follows: 1 min - 94°C / 2 sec - 94°C, 3 min - 72°C (7 cycles) / 2 sec - 94°C, 3 min - 67°C (32 cycles) / 5 min - 67°C.

The product of the first PCR reaction is diluted and used as a template for a second PCR reaction according to the manufacturer's instructions using a pair of nested primers which are located internally on the amplicon resulting from the first PCR reaction. For example, 5 μl of the reaction product of the first PCR reaction mixture may be diluted 180 times. Reactions are made in a 50 μl volume having a composition identical to that of the first PCR reaction except the nested primers are used. The first nested primer is specific for the adaptor, and is provided with the GenomeWalker™ kit. The second nested primer is specific for the particular extended cDNA or 5′ EST for which the promoter is to be cloned and should have a melting temperature, length, and location in the extended cDNA or 5′ EST which is consistent with its use in PCR reactions. The reaction parameters of the second PCR reaction are as follows: 1 min -

94°C / 2 sec - 94°C, 3 min - 72°C (6 cycles) / 2 sec - 94°C, 3 min - 67°C (25 cycles) / 5 min - 67°C. The product of the second PCR reaction is purified, cloned, and sequenced using standard techniques.

Alternatively, two or more human genomic DNA libraries can be constructed by using two or more restriction enzymes. The digested genomic DNA is cloned into vectors which can be converted into single stranded, circular, or linear DNA. A biotinylated oligonucleotide comprising at least 15 nucleotides from the extended cDNA or 5′ EST sequence is hybridized to the single stranded DNA. Hybrids between the biotinylated oligonucleotide and the single stranded DNA containing the extended cDNA or EST sequence are isolated as described in Example 29 above. Thereafter, the single stranded DNA containing the extended cDNA or EST sequence is released from the beads and converted into double stranded DNA using a primer specific for the extended cDNA or 5′ EST sequence or a primer corresponding to a sequence included in the cloning vector. The resulting double stranded DNA is transformed into bacteria. DNAs containing the 5′ EST or extended cDNA sequences are identified by colony PCR or colony hybridization.

Once the upstream genomic sequences have been cloned and sequenced as described above, prospective promoters and transcription start sites within the upstream sequences may be identified by comparing the sequences upstream of the extended cDNAs or 5′ ESTs with databases containing known transcription start sites, transcription factor binding sites, or promoter sequences.

In addition, promoters in the upstream sequences may be identified using promoter reporter vectors as described in Example .

EXAMPLE 59

Identification of Promoters in Cloned Upstream Sequences

The genomic sequences upstream of the extended cDNAs or 5′ ESTs are cloned into a suitable promoter reporter vector, such as the pSEAP-Basic, pSEAP-Enhancer, pβgal-Basic, pβgal-Enhancer, or pEGFP-1 Promoter Reporter vectors available from Clontech. Briefly, each of these promoter reporter vectors include multiple cloning sites positioned upstream of a reporter gene encoding a readily assayable protein such as secreted alkaline

phosphatase, β galactosidase, or green fluorescent protein. The sequences upstream of the extended cDNAs or 5' ESTs are inserted into the cloning sites upstream of the reporter gene in both orientations and introduced into an appropriate host cell. The level of reporter protein is assayed and compared to the level obtained from a vector which lacks an insert in the cloning site. The presence of an elevated expression level in the vector containing the insert with respect to the control vector indicates the presence of a promoter in the insert. If necessary, the upstream sequences can be cloned into vectors which contain an enhancer for augmenting transcription levels from weak promoter sequences. A significant level of expression above that observed with the vector lacking an insert indicates that a promoter sequence is present in the inserted upstream sequence.

Appropriate host cells for the promoter reporter vectors may be chosen based on the results of the above described determination of expression patterns of the extended cDNAs and ESTs. For example, if the expression pattern analysis indicates that the mRNA corresponding to a particular extended cDNA or 5' EST is expressed in fibroblasts, the promoter reporter vector may be introduced into a human fibroblast cell line.

Promoter sequences within the upstream genomic DNA may be further defined by constructing nested deletions in the upstream DNA using conventional techniques such as Exonuclease III digestion. The resulting deletion fragments can be inserted into the promoter reporter vector to determine whether the deletion has reduced or obliterated promoter activity. In this way, the boundaries of the promoters may be defined. If desired, potential individual regulatory sites within the promoter may be identified using site directed mutagenesis or linker scanning to obliterate potential transcription factor binding sites within the promoter individually or in combination. The effects of these mutations on transcription levels may be determined by inserting the mutations into the cloning sites in the promoter reporter vectors.

EXAMPLE 60

Cloning and Identification of Promoters

Using the method described in Example 58 above with 5' ESTs, sequences upstream of several genes were obtained. Using the primer pairs GGG AAG ATG GAG ATA GTA TTG CCT G (SEQ ID NO:29) and CTG CCA TGT ACA TGA TAG AGA GAT TC (SEQ

ID NO:30), the promoter having the internal designation P13H2 (SEQ ID NO:31) was obtained.

Using the primer pairs GTA CCA GGGG ACT GTG ACC ATT GC (SEQ ID NO:32) and CTG TGA CCA TTG CTC CCA AGA GAG (SEQ ID NO:33), the promoter having the internal designation P15B4 (SEQ ID NO:34) was obtained.

Using the primer pairs CTG GGA TGG AAG GCA CGG TA (SEQ ID NO:35) and GAG ACC ACA CAG CTA GAC AA (SEQ ID NO:36), the promoter having the internal designation P29B6 (SEQ ID NO:37) was obtained.

Figure 4 provides a schematic description of the promoters isolated and the way they are assembled with the corresponding 5' tags. The upstream sequences were screened for the presence of motifs resembling transcription factor binding sites or known transcription start sites using the computer program MatInspector release 2.0, August 1996.

Table VII describes the transcription factor binding sites present in each of these promoters. The columns labeled matrice provides the name of the MatInspector matrix used. The column labeled position provides the 5' position of the promoter site. Numeration of the sequence starts from the transcription site as determined by matching the genomic sequence with the 5' EST sequence. The column labeled "orientation" indicates the DNA strand on which the site is found, with the + strand being the coding strand as determined by matching the genomic sequence with the sequence of the 5' EST. The column labeled "score" provides the MatInspector score found for this site. The column labeled "length" provides the length of the site in nucleotides. The column labeled "sequence" provides the sequence of the site found.

Bacterial clones containing plasmids containing the promoter sequences described above described above are presently stored in the inventor's laboratories under the internal identification numbers provided above. The inserts may be recovered from the deposited materials by growing an aliquot of the appropriate bacterial clone in the appropriate medium. The plasmid DNA can then be isolated using plasmid isolation procedures familiar to those skilled in the art such as alkaline lysis minipreps or large scale alkaline lysis plasmid isolation procedures. If desired the plasmid DNA may be further enriched by centrifugation on a cesium chloride gradient, size exclusion chromatography, or anion exchange chromatography. The plasmid DNA obtained using these procedures may then be manipulated using standard

cloning techniques familiar to those skilled in the art. Alternatively, a PCR can be done with primers designed at both ends of the EST insertion. The PCR product which corresponds to the 5′ EST can then be manipulated using standard cloning techniques familiar to those skilled in the art.

The promoters and other regulatory sequences located upstream of the extended cDNAs or 5′ ESTs may be used to design expression vectors capable of directing the expression of an inserted gene in a desired spatial, temporal, developmental, or quantitative manner. A promoter capable of directing the desired spatial, temporal, developmental, and quantitative patterns may be selected using the results of the expression analysis described in Example 26 above. For example, if a promoter which confers a high level of expression in muscle is desired, the promoter sequence upstream of an extended cDNA or 5′ EST derived from an mRNA which is expressed at a high level in muscle, as determined by the method of Example 26, may be used in the expression vector.

Preferably, the desired promoter is placed near multiple restriction sites to facilitate the cloning of the desired insert downstream of the promoter, such that the promoter is able to drive expression of the inserted gene. The promoter may be inserted in conventional nucleic acid backbones designed for extrachromosomal replication, integration into the host chromosomes or transient expression. Suitable backbones for the present expression vectors include retroviral backbones, backbones from eukaryotic episomes such as SV40 or Bovine Papilloma Virus, backbones from bacterial episomes, or artificial chromosomes.

Preferably, the expression vectors also include a polyA signal downstream of the multiple restriction sites for directing the polyadenylation of mRNA transcribed from the gene inserted into the expression vector.

Following the identification of promoter sequences using the procedures of Examples 58-60, proteins which interact with the promoter may be identified as described in Example 61 below.

EXAMPLE 61

Identification of Proteins Which Interact with Promoter Sequences, Upstream Regulatory Sequences, or mRNA

Sequences within the promoter region which are likely to bind transcription factors may be identified by homology to known transcription factor binding sites or through conventional mutagenesis or deletion analyses of reporter plasmids containing the promoter sequence. For example, deletions may be made in a reporter plasmid containing the promoter sequence of interest operably linked to an assayable reporter gene. The reporter plasmids carrying various deletions within the promoter region are transfected into an appropriate host cell and the effects of the deletions on expression levels is assessed. Transcription factor binding sites within the regions in which deletions reduce expression levels may be further localized using site directed mutagenesis, linker scanning analysis, or other techniques familiar to those skilled in the art.

Nucleic acids encoding proteins which interact with sequences in the promoter may be identified using one-hybrid systems such as those described in the manual accompanying the Matchmaker One-Hybrid System kit available from Clontech (Catalog No. K1603-1), the disclosure of which is incorporated herein by reference. Briefly, the Matchmaker One-hybrid system is used as follows. The target sequence for which it is desired to identify binding proteins is cloned upstream of a selectable reporter gene and integrated into the yeast genome. Preferably, multiple copies of the target sequences are inserted into the reporter plasmid in tandem. A library comprised of fusions between cDNAs to be evaluated for the ability to bind to the promoter and the activation domain of a yeast transcription factor, such as GAL4, is transformed into the yeast strain containing the integrated reporter sequence. The yeast are plated on selective media to select cells expressing the selectable marker linked to the promoter sequence. The colonies which grow on the selective media contain genes encoding proteins which bind the target sequence. The inserts in the genes encoding the fusion proteins are further characterized by sequencing. In addition, the inserts may be inserted into expression vectors or *in vitro* transcription vectors. Binding of the polypeptides encoded by the inserts to the promoter DNA may be confirmed by techniques familiar to those skilled in the art, such as gel shift analysis or DNAse protection analysis.

VII. Use of 5′ ESTs (or cDNAs or Genomic DNAs Obtainable Therefrom) in Gene Therapy

The present invention also comprises the use of 5'ESTs (or cDNA or genomic DNA obtainable therefrom) in gene therapy strategies, including antisense and triple helix strategies as described in Examples 62 and 63 below. In antisense approaches, nucleic acid sequences complementary to an mRNA are hybridized to the mRNA intracellularly, thereby blocking the expression of the protein encoded by the mRNA. The antisense sequences may prevent gene expression through a variety of mechanisms. For example, the antisense sequences may inhibit the ability of ribosomes to translate the mRNA. Alternatively, the antisense sequences may block transport of the mRNA from the nucleus to the cytoplasm, thereby limiting the amount of mRNA available for translation. Another mechanism through which antisense sequences may inhibit gene expression is by interfering with mRNA splicing. In yet another strategy, the antisense nucleic acid may be incorporated in a ribozyme capable of specifically cleaving the target mRNA.

EXAMPLE 62

Preparation and Use of Antisense Oligonucleotides

The antisense nucleic acid molecules to be used in gene therapy may be either DNA or RNA sequences. They may comprise a sequence complementary to the sequence of the 5'EST (or cDNA or genomic DNA obtainable therefrom). The antisense nucleic acids should have a length and melting temperature sufficient to permit formation of an intracellular duplex with sufficient stability to inhibit the expression of the mRNA in the duplex. Strategies for designing antisense nucleic acids suitable for use in gene therapy are disclosed in Green *et al.*, *Ann. Rev. Biochem.* 55:569-597, 1986; and Izant and Weintraub, *Cell* 36:1007-1015, 1984, which are hereby incorporated by reference.

In some strategies, antisense molecules are obtained from a nucleotide sequence encoding a protein by reversing the orientation of the coding region with respect to a promoter so as to transcribe the opposite strand from that which is normally transcribed in the cell. The antisense molecules may be transcribed using *in vitro* transcription systems such as those which employ T7 or SP6 polymerase to generate the transcript. Another approach

involves transcription of the antisense nucleic acids *in vivo* by operably linking DNA containing the antisense sequence to a promoter in an expression vector.

Alternatively, oligonucleotides which are complementary to the strand normally transcribed in the cell may be synthesized *in vitro*. Thus, the antisense nucleic acids are complementary to the corresponding mRNA and are capable of hybridizing to the mRNA to create a duplex. In some embodiments, the antisense sequences may contain modified sugar phosphate backbones to increase stability and make them less sensitive to RNase activity. Examples of modifications suitable for use in antisense strategies are described by Rossi *et al.*, *Pharmacol. Ther.* **50(2)**:245-254, 1991, which is hereby incorporated by reference.

Various types of antisense oligonucleotides complementary to the sequence of the 5'EST (or cDNA or genomic DNA obtainable therefrom) may be used. In one preferred embodiment, stable and semi-stable antisense oligonucleotides described in International Application No. PCT WO94/23026, hereby incorporated by reference, are used. In these molecules, the 3′ end or both the 3′ and 5′ ends are engaged in intramolecular hydrogen bonding between complementary base pairs. These molecules are better able to withstand exonuclease attacks and exhibit increased stability compared to conventional antisense oligonucleotides.

In another preferred embodiment, the antisense oligodeoxynucleotides against herpes simplex virus types 1 and 2 described in International Application No. WO 95/04141, hereby incorporated by reference, are used.

In yet another preferred embodiment, the covalently cross-linked antisense oligonucleotides described in International Application No. WO 96/31523, hereby incorporated by reference, are used. These double- or single-stranded oligonucleotides comprise one or more, respectively, inter- or intra-oligonucleotide covalent cross-linkages, wherein the linkage consists of an amide bond between a primary amine group of one strand and a carboxyl group of the other strand or of the same strand, respectively, the primary amine group being directly substituted in the 2' position of the strand nucleotide monosaccharide ring, and the carboxyl group being carried by an aliphatic spacer group substituted on a nucleotide or nucleotide analog of the other strand or the same strand, respectively.

The antisense oligodeoxynucleotides and oligonucleotides disclosed in International Application No. WO 92/18522, incorporated by reference, may also be used. These molecules are stable to degradation and contain at least one transcription control recognition sequence which binds to control proteins and are effective as decoys therefore. These molecules may contain "hairpin" structures, "dumbbell" structures, "modified dumbbell" structures, "cross-linked" decoy structures and "loop" structures.

In another preferred embodiment, the cyclic double-stranded oligonucleotides described in European Patent Application No. 0 572 287 A2, hereby incorporated by reference are used. These ligated oligonucleotide "dumbbells" contain the binding site for a transcription factor and inhibit expression of the gene under control of the transcription factor by sequestering the factor.

Use of the closed antisense oligonucleotides disclosed in International Application No. WO 92/19732, hereby incorporated by reference, is also contemplated. Because these molecules have no free ends, they are more resistant to degradation by exonucleases than are conventional oligonucleotides. These oligonucleotides may be multifunctional, interacting with several regions which are not adjacent to the target mRNA.

The appropriate level of antisense nucleic acids required to inhibit gene expression may be determined using *in vitro* expression analysis. The antisense molecule may be introduced into the cells by diffusion, injection, infection, transfection or h-region-mediated import using procedures known in the art. For example, the antisense nucleic acids can be introduced into the body as a bare or naked oligonucleotide, oligonucleotide encapsulated in lipid, oligonucleotide sequence encapsidated by viral protein, or as an oligonucleotide operably linked to a promoter contained in an expression vector. The expression vector may be any of a variety of expression vectors known in the art, including retroviral or viral vectors, vectors capable of extrachromosomal replication, or integrating vectors. The vectors may be DNA or RNA.

The antisense molecules are introduced onto cell samples at a number of different concentrations preferably between $1x10^{-10}$M to $1x10^{-4}$M. Once the minimum concentration that can adequately control gene expression is identified, the optimized dose is translated into a dosage suitable for use *in vivo*. For example, an inhibiting concentration in culture of $1x10^{-7}$ translates into a dose of approximately 0.6 mg/kg bodyweight. Levels of oligonucleotide

approaching 100 mg/kg bodyweight or higher may be possible after testing the toxicity of the oligonucleotide in laboratory animals. It is additionally contemplated that cells from the vertebrate are removed, treated with the antisense oligonucleotide, and reintroduced into the vertebrate.

It is further contemplated that the antisense oligonucleotide sequence is incorporated into a ribozyme sequence to enable the antisense to specifically bind and cleave its target mRNA. For technical applications of ribozyme and antisense oligonucleotides see Rossi *et al.*, *supra*.

In a preferred application of this invention, the polypeptide encoded by the gene is first identified, so that the effectiveness of antisense inhibition on translation can be monitored using techniques that include but are not limited to antibody-mediated tests such as RIAs and ELISA, functional assays, or radiolabeling.

The 5′ ESTs of the present invention (or cDNAs or genomic DNAs obtainable therefrom) may also be used in gene therapy approaches based on intracellular triple helix formation. Triple helix oligonucleotides are used to inhibit transcription from a genome. They are particularly useful for studying alterations in cell activity as it is associated with a particular gene. The 5′ EST sequences (or cDNAs or genomic DNAs obtainable therefrom) of the present invention or, more preferably, a portion of those sequences, can be used to inhibit gene expression in individuals having diseases associated with expression of a particular gene. Similarly, a portion of 5′ EST sequences (or cDNAs or genomic DNAs obtainable therefrom) can be used to study the effect of inhibiting transcription of a particular gene within a cell. Traditionally, homopurine sequences were considered the most useful for triple helix strategies. However, homopyrimidine sequences can also inhibit gene expression. Such homopyrimidine oligonucleotides bind to the major groove at homopurine:homopyrimidine sequences. Thus, both types of sequences from the 5'EST or from the gene corresponding to the 5'EST are contemplated within the scope of this invention.

EXAMPLE 63

Preparation and Use of Triple Helix Probes

The sequences of the 5′ ESTs (or cDNAs or genomic DNAs obtainable therefrom) are scanned to identify 10-mer to 20-mer homopyrimidine or homopurine stretches which could be used in triple-helix based strategies for inhibiting gene expression. Following identification of candidate homopyrimidine or homopurine stretches, their efficiency in inhibiting gene expression is assessed by introducing varying amounts of oligonucleotides containing the candidate sequences into tissue culture cells which normally express the target gene. The oligonucleotides may be prepared on an oligonucleotide synthesizer or they may be purchased commercially from a company specializing in custom oligonucleotide synthesis, such as GENSET, Paris, France.

The oligonucleotides may be introduced into the cells using a variety of methods known to those skilled in the art, including but not limited to calcium phosphate precipitation, DEAE-Dextran, electroporation, liposome-mediated transfection or native uptake.

Treated cells are monitored for altered cell function or reduced gene expression using techniques such as Northern blotting, RNase protection assays, or PCR based strategies to monitor the transcription levels of the target gene in cells which have been treated with the oligonucleotide. The cell functions to be monitored are predicted based upon the homologies of the target gene corresponding to the extended cDNA from which the oligonucleotide was derived with known gene sequences that have been associated with a particular function. The cell functions can also be predicted based on the presence of abnormal physiologies within cells derived from individuals with a particular inherited disease, particularly when the extended cDNA is associated with the disease using techniques described in Example 56.

The oligonucleotides which are effective in inhibiting gene expression in tissue culture cells may then be introduced *in vivo* using the techniques described above and in Example 62 at a dosage calculated based on the *in vitro* results, as described in Example 62.

In some embodiments, the natural (beta) anomers of the oligonucleotide units can be replaced with alpha anomers to render the oligonucleotide more resistant to nucleases. Further, an intercalating agent such as ethidium bromide, or the like, can be attached to the 3′ end of the alpha oligonucleotide to stabilize the triple helix. For information on the

generation of oligonucleotides suitable for triple helix formation see Griffin *et al.*, *Science* 245:967-971, 1989, which is hereby incorporated by this reference.

EXAMPLE 64

Use of cDNAs Obtained Using the 5′ ESTs to Express an Encoded Protein in a Host Organism

The cDNAs obtained as described above using the 5′ ESTs of the present invention may also be used to express an encoded protein in a host organism to produce a beneficial effect. In such procedures, the encoded protein may be transiently expressed in the host organism or stably expressed in the host organism. The encoded protein may have any of the activities described above. The encoded protein may be a protein which the host organism lacks or, alternatively, the encoded protein may augment the existing levels of the protein in the host organism.

A full length extended cDNA encoding the signal peptide and the mature protein, or an extended cDNA encoding only the mature protein is introduced into the host organism. The extended cDNA may be introduced into the host organism using a variety of techniques known to those of skill in the art. For example, the extended cDNA may be injected into the host organism as naked DNA such that the encoded protein is expressed in the host organism, thereby producing a beneficial effect.

Alternatively, the extended cDNA may be cloned into an expression vector downstream of a promoter which is active in the host organism. The expression vector may be any of the expression vectors designed for use in gene therapy, including viral or retroviral vectors. The expression vector may be directly introduced into the host organism such that the encoded protein is expressed in the host organism to produce a beneficial effect. In another approach, the expression vector may be introduced into cells *in vitro*. Cells containing the expression vector are thereafter selected and introduced into the host organism, where they express the encoded protein to produce a beneficial effect.

EXAMPLE 65

Use of Signal Peptides Encoded by 5' ESTs or Sequences obtained Therefrom to Import Proteins Into Cells

The short core hydrophobic region (h) of signal peptides encoded by the 5'ESTS or extended cDNAs derived from SEQ ID NOs: 38-195 may also be used as a carrier to import a peptide or a protein of interest, so-called cargo, into tissue culture cells (Lin *et al., J. Biol. Chem.*, **270**: 14225-14258, 1995; Du *et al., J. Peptide Res.*, **51**: 235-243, 1998; Rojas *et al., Nature Biotech.*, **16**: 370-375, 1998).

When cell permeable peptides of limited size (approximately up to 25 amino acids) are to be translocated across cell membrane, chemical synthesis may be used in order to add the h region to either the C-terminus or the N-terminus to the cargo peptide of interest. Alternatively, when longer peptides or proteins are to be imported into cells, nucleic acids can be genetically engineered, using techniques familiar to those skilled in the art, in order to link the extended cDNA sequence encoding the h region to the 5' or the 3' end of a DNA sequence coding for a cargo polypeptide. Such genetically engineered nucleic acids are then translated either *in vitro* or *in vivo* after transfection into appropriate cells, using conventional techniques to produce the resulting cell permeable polypeptide. Suitable hosts cells are then simply incubated with the cell permeable polypeptide which is then translocated across the membrane.

This method may be applied to study diverse intracellular functions and cellular processes. For instance, it has been used to probe functionally relevant domains of intracellular proteins and to examine protein-protein interactions involved in signal transduction pathways (Lin *et al., supra*; Lin *et al., J. Biol. Chem.*, **271**: 5305-5308, 1996; Rojas *et al., J. Biol. Chem.*, **271**: 27456-27461, 1996; Liu *et al., Proc. Natl. Acad. Sci. USA*, **93**: 11819-11824, 1996; Rojas *et al., Bioch. Biophys. Res. Commun.*, **234**: 675-680, 1997).

Such techniques may be used in cellular therapy to import proteins producing therapeutic effects. For instance, cells isolated from a patient may be treated with imported therapeutic proteins and then re-introduced into the host organism.

Alternatively, the h region of signal peptides of the present invention could be used in combination with a nuclear localization signal to deliver nucleic acids into cell nucleus. Such oligonucleotides may be antisense oligonucleotides or oligonucleotides designed to form

triple helixes, as described in examples 62 and 63 respectively, in order to inhibit processing and/or maturation of a target cellular RNA.

As discussed above, the cDNAs or portions thereof obtained using the 5′ ESTs of the present invention can be used for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination for expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris *et al.*, *Cell* 75:791-803, 1993, the disclosure of which is hereby incorporated by reference) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The proteins or polypeptides provided by the present invention can similarly be used in assays to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins

involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation *Molecular Cloning; A Laboratory Manual*, 2d ed., Cold Spring Harbor Laboratory Press, Sambrook,. Fritsch and Maniatis eds., 1989, and *Methods in Enzymology; Guide to Molecular Cloning Techniques*, Academic Press, Berger and Kimmel eds., 1987.

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

Although this invention has been described in terms of certain preferred embodiments, other embodiments which will be apparent to those of ordinary skill in the art in view of the disclosure herein are also within the scope of this invention. Accordingly, the scope of the invention is intended to be defined only by reference to the appended claims. All documents cited herein are incorporated herein by reference in their entirety.

	Search characteristic		Selection Characteristics		
Step	Program	Strand	Parameters	Identity (%)	Length (bp)
miscellanaeous	blastn	both	S=61 X=16	90	17
tRNA	fasta	both	-	80	60
rRNA	blastn	both	S=108	80	40
mtRNA	blastn	both	S=108	80	40
Procaryotic	blastn	both	S=144	90	40
Fungal	blastn	both	S=144	90	40
Alu	fasta*	both	-	70	40
L1	blastn	both	S=72	70	40
Repeats	blastn	both	S=72	70	40
Promoters	blastn	top	S=54 X=16	90	15†
Vertebrate	fasta*	both	S=108	90	30
ESTs	blastn	both	S=108 X=16	90	30
Proteins	blastx¤	top	E = 0.001	-	-

Table 1: Parameters used for each step of EST analysis

* use "Quick Fast" Database scanner

† alignement further constrained to begin closer than 10bp to EST\5' end

¤ using BLOSUM62 substitution matrix

TABLE II

SEQ. ID NO.	CATEGORY	VON HEIJNE SCORE	TISSUE SOURCE	INTERNAL DESIGNATION
ID38	new	14.3	Substantia nigra	47-39-4-A10-PU
ID39	new	11.1	Fetal brain	57-9-4-C5-PU
ID40	new	10.6	Fetal brain	57-19-1-B11-PU
ID41	new	9.1	Fetal brain	57-7-1-G12-PU
ID42	new	8.8	Substantia nigra	47-22-3-D2-PU
ID43	new	8.7	Fetal brain	57-21-2-H11-PU
ID44	new	8.4	Substantia nigra	47-37-3-F6-PU
ID45	new	8.2	Substantia nigra	47-54-1-A8-PU
ID46	new	8.2	Substantia nigra	47-15-1-E5-PU
ID47	new	8	Substantia nigra	47-24-1-A6-PU
ID48	new	7.8	Fetal brain	57-10-3-H10-PU
ID49	new	7.7	Substantia nigra	47-17-1-D7-PU
ID50	new	7.6	Cerebellum	55-9-4-A4-PU
ID51	new	7.5	Substantia nigra	47-18-3-C2-PU
ID52	new	7.4	Fetal brain	57-19-1-C8-PU
ID53	new	7.4	Substantia nigra	47-1-4-C5-PU
ID54	new	7.2	Substantia nigra	47-24-1-B5-PU
ID55	new	7.2	Cerebellum	55-6-2-A9-PU
ID56	new	7.1	Substantia nigra	47-30-4-A8-PU
ID57	new	6.9	Fetal brain	57-21-1-D5-PU
ID58	new	6.9	Substantia nigra	47-4-2-C7-PU
ID59	new	6.8	Substantia nigra	47-12-3-A8-PU
ID60	new	6.8	Substantia nigra	47-7-4-E6-PU
ID61	new	6.8	Substantia nigra	47-15-3-G3-PU
ID62	new	6.7	Fetal brain	57-20-2-B5-PU
ID63	new	6.6	Substantia nigra	47-2-2-E6-PU
ID64	new	6.5	Fetal brain	57-25-1-G3-PU
ID65	new	6.5	Fetal brain	57-7-4-B12-PU
ID66	new	6.5	Substantia nigra	47-21-1-D9-PU
ID67	new	6.4	Substantia nigra	47-31-2-H9-PU
ID68	new	5.7	Cerebellum	55-1-3-D11-PU
ID69	new	5.7	Cerebellum	55-7-2-A1-PU
ID70	new	5.7	Fetal brain	57-28-3-C1-PU
ID71	new	5.6	Fetal brain	57-9-4-D11-PU
ID72	new	5.5	Substantia nigra	47-7-4-C10-PU
ID73	new	5.5	Fetal brain	57-22-1-E11-PU
ID74	new	5.4	Fetal brain	57-20-2-D9-PU
ID75	new	5.4	Substantia nigra	47-39-3-E7-PU
ID76	new	5.4	Surrenals	62-3-1-G5-PU
ID77	new	5.4	Fetal brain	57-18-4-H5-PU
ID78	new	5.4	Fetal brain	57-22-2-H8-PU
ID79	new	5.3	Fetal brain	57-22-2-E12-PU
ID80	new	5.3	Fetal brain	57-23-3-B8-PU
ID81	new	5.2	Fetal brain	57-6-3-C5-PU
ID82	new	5.2	Substantia nigra	47-7-1-D2-PU
ID83	new	5.1	Fetal brain	57-7-2-G9-PU
ID84	new	5.1	Fetal brain	57-10-3-D3-PU
ID85	new	5.1	Substantia nigra	47-4-4-F2-PU
ID86	new	5.1	Fetal brain	57-4-4-H6-PU

SEQ. ID NO.	CATEGORY	VON HEIJNE SCORE	TISSUE SOURCE	INTERNAL DESIGNATION
ID87	new	5	Substantia nigra	47-10-2-G12-PU
ID88	new	5	Cerebellum	55-10-3-E12-PU
ID89	new	4.9	Substantia nigra	47-8-2-D1-PU
ID90	new	4.9	Fetal brain	57-3-4-C9-PU
ID91	new	4.9	Substantia nigra	47-14-1-C3-PU
ID92	new	4.8	Substantia nigra	47-3-4-C8-PU
ID93	new	4.8	Substantia nigra	47-15-1-B10-PU
ID94	new	4.8	Fetal brain	57-26-3-A12-PU
ID95	new	4.7	Substantia nigra	47-26-3-B10-PU
ID96	new	4.7	Substantia nigra	47-26-1-B6-PU
ID97	new	4.7	Surrenals	62-5-1-B8-PU
ID98	new	4.6	Substantia nigra	47-15-4-H9-PU
ID99	new	4.6	Cerebellum	55-2-4-D3-PU
ID100	new	4.5	Fetal brain	57-6-1-B1-PU
ID101	new	4.5	Fetal brain	57-26-4-E4-PU
ID102	new	4.5	Substantia nigra	47-2-2-A7-PU
ID103	new	4.5	Substantia nigra	47-55-2-B3-PU
ID104	new	4.5	Substantia nigra	47-54-1-C9-PU
ID105	new	4.4	Cerebellum	55-8-2-A2-PU
ID106	new	4.4	Substantia nigra	47-4-2-H4-PU
ID107	new	4.3	Fetal brain	57-27-3-B11-PU
ID108	new	4.2	Fetal brain	57-22-4-D2-PU
ID109	new	4.2	Substantia nigra	47-20-4-E2-PU
ID110	new	4.2	Substantia nigra	47-2-3-H2-PU
ID111	new	4.1	Substantia nigra	47-22-3-G5-PU
ID112	new	4.1	Fetal brain	57-18-3-A5-PU
ID113	new	4.1	Fetal brain	57-9-3-H7-PU
ID114	new	4.1	Surrenals	62-5-2-B6-PU
ID115	new	4	Cerebellum	55-12-1-E12-PU
ID116	new	4	Fetal brain	57-20-1-A5-PU
ID117	new	3.9	Substantia nigra	47-22-4-F6-PU
ID118	new	3.8	Fetal brain	57-19-3-E1-PU
ID119	new	3.8	Substantia nigra	47-18-3-G5-PU
ID120	new	3.8	Substantia nigra	47-20-1-G3-PU
ID121	new	3.8	Fetal brain	57-6-4-A1-PU
ID122	new	3.8	Fetal brain	57-27-3-G10-PU
ID123	new	3.7	Substantia nigra	47-2-4-C7-PU
ID124	new	3.7	Cerebellum	55-6-1-E6-PU
ID125	new	3.6	Fetal brain	57-4-4-F7-PU
ID126	new	3.6	Substantia nigra	47-30-2-B1-PU
ID127	new	3.6	Substantia nigra	47-29-1-F11-PU
ID128	new	3.6	Substantia nigra	47-39-3-D4-PU
ID129	new	3.5	Substantia nigra	47-15-2-G3-PU
ID130	new	3.5	Fetal brain	57-18-3-E6-PU
ID131	new	3.5	Substantia nigra	47-40-2-G6-PU
ID132	new	3.5	Fetal brain	57-6-4-D7-PU
ID133	new	3.5	Substantia nigra	47-55-4-A8-PU
ID134	ext-est-not-vrt	9.8	Substantia nigra	47-39-4-B9-PU
ID135	ext-est-not-vrt	9.2	Cerebellum	55-11-1-H5-PU
ID136	ext-est-not-vrt	9	Substantia nigra	47-4-4-G1-PU
ID137	ext-est-not-vrt	7.2	Substantia nigra	47-2-3-G9-PU

SEQ. ID NO.	CATEGORY	VON HEIJNE SCORE	TISSUE SOURCE	INTERNAL DESIGNATION
ID138	ext-est-not-vrt	7.2	Cerebellum	55-10-3-F5-PU
ID139	ext-est-not-vrt	5.6	Fetal brain	57-4-4-G6-PU
ID140	ext-est-not-vrt	4.2	Cerebellum	55-7-1-D11-PU
ID141	ext-est-not-vrt	3.7	Substantia nigra	47-19-2-F7-PU
ID142	ext-est-not-vrt	3.7	Substantia nigra	47-1-4-D2-PU
ID143	ext-est-not-vrt	3.6	Cerebellum	55-5-4-A6-PU
ID144	ext-est-not-vrt	3.6	Cerebellum	55-4-4-H3-PU
ID145	ext-est-not-vrt	3.5	Cerebellum	55-3-1-G6-PU
ID146	ext-est-not-vrt	3.5	Substantia nigra	47-55-2-H2-PU
ID147	est-not-ext	12.4	Substantia nigra	47-39-4-H8-PU
ID148	est-not-ext	11.4	Fetal brain	57-26-4-A4-PU
ID149	est-not-ext	11.1	Substantia nigra	47-2-3-D1-PU
ID150	est-not-ext	9.2	Substantia nigra	47-4-1-E4-PU
ID151	est-not-ext	9	Substantia nigra	47-40-4-G9-PU
ID152	est-not-ext	8.8	Fetal brain	57-5-4-G3-PU
ID153	est-not-ext	7.5	Substantia nigra	47-13-4-C1-PU
ID154	est-not-ext	7.4	Fetal brain	57-20-4-E2-PU
ID155	est-not-ext	7	Substantia nigra	47-24-4-H4-PU
ID156	est-not-ext	6.9	Substantia nigra	47-26-2-B2-PU
ID157	est-not-ext	6.8	Substantia nigra	47-11-1-A2-PU
ID158	est-not-ext	6.4	Fetal brain	57-19-2-G8-PU
ID159	est-not-ext	6.4	Fetal brain	57-19-4-H8-PU
ID160	est-not-ext	6.3	Substantia nigra	47-39-2-A11-PU
ID161	est-not-ext	6.2	Fetal brain	57-24-2-B4-PU
ID162	est-not-ext	5.9	Surrenals	62-1-1-G3-PU
ID163	est-not-ext	5.7	Fetal brain	57-2-4-H4-PU
ID164	est-not-ext	5.6	Fetal brain	57-8-2-D3-PU
ID165	est-not-ext	5.5	Cerebellum	55-11-4-G2-PU
ID166	est-not-ext	5.4	Substantia nigra	47-24-1-B6-PU
ID167	est-not-ext	5.4	Substantia nigra	47-55-3-B10-PU
ID168	est-not-ext	5.4	Surrenals	62-8-1-B12-PU
ID169	est-not-ext	5.3	Substantia nigra	47-39-1-C9-PU
ID170	est-not-ext	5.3	Fetal brain	57-20-2-F1-PU
ID171	est-not-ext	5.2	Fetal brain	57-25-1-F10-PU
ID172	est-not-ext	5.2	Fetal brain	57-28-4-B12-PU
ID173	est-not-ext	5.1	Substantia nigra	47-15-2-D12-PU
ID174	est-not-ext	5.1	Substantia nigra	47-2-3-G3-PU
ID175	est-not-ext	4.9	Substantia nigra	47-40-3-D8-PU
ID176	est-not-ext	4.9	Substantia nigra	47-40-3-G11-PU
ID177	est-not-ext	4.9	Substantia nigra	47-14-3-D2-PU
ID178	est-not-ext	4.8	Substantia nigra	47-19-1-B7-PU
ID179	est-not-ext	4.8	Substantia nigra	47-19-1-A3-PU
ID180	est-not-ext	4.7	Substantia nigra	47-55-3-G2-PU
ID181	est-not-ext	4.6	Substantia nigra	47-3-4-G7-PU
ID182	est-not-ext	4.5	Substantia nigra	47-29-1-B7-PU
ID183	est-not-ext	4.4	Fetal brain	57-21-4-G6-PU
ID184	est-not-ext	4.3	Substantia nigra	47-2-1-E12-PU
ID185	est-not-ext	4.3	Substantia nigra	47-9-4-D2-PU
ID186	est-not-ext	4.3	Fetal brain	57-2-4-F8-PU
ID187	est-not-ext	4.3	Fetal brain	57-18-1-D5-PU
ID188	est-not-ext	4.2	Substantia nigra	47-8-4-D2-PU

SEQ. ID NO.	CATEGORY	VON HEIJNE SCORE	TISSUE SOURCE	INTERNAL DESIGNATION
ID189	est-not-ext	4.1	Substantia nigra	47-17-3-H11-PU
ID190	est-not-ext	3.9	Fetal brain	57-28-2-G6-PU
ID191	est-not-ext	3.7	Fetal brain	57-27-3-G1-PU
ID192	est-not-ext	3.7	Substantia nigra	47-37-4-G11-PU
ID193	est-not-ext	3.7	Surrenals	62-11-3-A2-PU
ID194	ext-vrt-not-genomic	10.9	Surrenals	62-10-2-E4-PU
ID195	ext-vrt-not-genomic	8.9	Substantia nigra	47-14-3-H7-PU

TABLE III

SEQ. ID NO.	SIGNAL PEPTIDE
ID38	MRVRIGLTLLLCAVLLSLASA
ID39	MGLHLRPYRVGLLPDGLLFLLLLLMLLA
ID40	MATLSFVFLLLGAVSWPPASA
ID41	MFLFLSPATPVLPPSLDSRDLLPHLFWGRAGSSSSSPALSPVLCLRGLVSLAFQ
ID42	MWAMESGHLLWALLFMQSLWP
ID43	MTHYRNILGLLCCVLATMA
ID44	MKLLLLLASLIERSS
ID45	MARNQALVCLPSFQNAFIPVEDLPTSFXLFLALCASFS
ID46	MPNESWQIPCGKQEAETLFNFQSLLLLFYSFYVLA
ID47	MQTTFIDVTVDQHVAKSNDHLSVLVLLICLVSSYLP
ID48	MALGEEKAEAEASXDTKAQSYGRGSCRERELDIPGPMSGEQPPRLEAEGGLISPVWGAEX YLPLLAGLGLTLA
ID49	MTSLYLKHLLCISPFVPFTSG
ID50	MTDSPNAHGLALTTKWMMPAVSLNLTYYLPSWYLCLATLTLFHTSFS
ID51	MASSHWNETTTSVYQYLGFQVQKIYPFHDNWNTACFVILLLFIFTVVS
ID52	MSLLFVFCLECSIFLLNMWVACLLS
ID53	MFVVTVLLLLPLVAFITL
ID54	MNRSCRNTGIIYALQFLFLVFA
ID55	MTQTTWGAPTRASNHPLPAWLTLSLLLAWVTLTHL
ID56	MLKXXAVLCVCAAAWC
ID57	MISAHCNLHLLGSSISPASA
ID58	MXXKACRTLAWLPXPFLPFLLSLPLDQT
ID59	MAVKRLGLLLVFLPHPQRG
ID60	MQAVDNLTSAPGNTSLCTRDYKITQVLFPLLYTVLFFVGLITNGLAMRIFFQIRSKSNFI IFLKNTVISDLLMILTFPFKILS
ID61	MWTLPSLSASFQPFLGSLRPSHILWFFLPSLXCPEC
ID62	MSLTDVPMSLLLFQPSSHSATG
ID63	MDWSLAFLLVSLYWSHM
ID64	MYLLILFFMVGRIIP
ID65	MNKPPWEESWGQNQLSGEPATWSLCISPLPGREPSLLVVSCCLLFHQA
ID66	MLILELTMMLSFLILLSIDSLVSG
ID67	MKLQRSRAFRIECSAILRRAERLVWNDVCSESQSQSRDSCLLGAAWASRLRT
ID68	MVIFTLCVFTLPFLCA
ID69	MWGALPVLVVGTWSSQGQA
ID70	MTRLVCGFLQISLSLA
ID71	MNFLLPLLLHHLTFH
ID72	MLSARDRRDRHPEEGVVAELQGFAVDKAFLTSHKGILLETELALTLIIFICXTASISA
ID73	MLTMSVTLSPLRS
ID74	MFXPVALIFPISVSDPTIHPITQAQNLESXLQSFFLLISSVRPISQ
ID75	MLLFFPFFGETVSLHHPCWCAVLRSWLAASS
ID76	MPLKNLFSVGLWDPYNLLKKHVLVVVCYLSWRVSS
ID77	MAMAQKLSHLLPSLRQVIQ
ID78	MIAPTLKGTPSSSAPLALVALAPHSVQK
ID79	MCLFPVSPCPAYSFSSEXXGAVLLLVESLCLVFNLLS
ID80	MKIAVLFCFLLLIIF
ID81	MCSPRSPLNLSLVPVGAVLLSSLPISP
ID82	MGLHISLIKFLLANGPHIPSHQRPFEPKGEKSCRIEVVTLPLTSHCLA
ID83	MKTTYVIFMQSKALLTLYVFVASSMQ
ID84	MNALVFLIFLRFINI
ID85	MQLGPLHTVSTPFFFCWGFLLTGHSLSHS
ID86	MGRGWERTVCSLGWRGGPDPLSWATCWSGARSRHTRVSSIVNGYVGSVCCCVGPLRG

SEQ. ID NO.	SIGNAL PEPTIDE
ID87	MPEAVEQSAHLFVTWSSQRALS
ID88	MPGTHTFTFKSCWLIALSVPLVFW
ID89	MLLLTFKWFLFCLIGLDLLCQV
ID90	MATTGRRQAEPPPVRPAHSRPPPRVPGSSSLGLAGLMSPVPNLHLLLPLTTP
ID91	MVPFIYLQAHFTLCSG
ID92	MFFSFLLTINLVSL
ID93	MWPGRECKNWGLLCFASECTT
ID94	MLTPFSLEEKLLECHYVLAKLAGACLLLTLRQPPTHS
ID95	MKRIKSMMGKVEHIKIKGEKQRSRHVKIVFVGLIFLKSSA
ID96	MQSALCLFCKICPFTHG
ID97	MMHNIIVKELIVTFFLGITVVQMLISVTGLKGVEAQNGSESEVFVGKYETLVFYWPSLLC LAFLLGRFL
ID98	MSNCLQNFLKITSTRLLCSRLCQQLRS
ID99	MSGXGLFLRTTAAARA
ID100	MNPLHKHCAAGPLTWLHLLLSHLKS
ID101	MPKDKRGARHNSPHFSFAVLRVLHLPALT
ID102	MTIHVLRKCCQMGRLNNEWLPGLVIPLCVSRQLLTGART
ID103	MQAASFGRGRNGLDNWGIAALLGLLQLRFK
ID104	MSPSLGDRCSSWLHLVSHLESISGPLLNIPENLLLCCHRCTNC
ID105	MSGAEPTTFIRYFLLPCLINLAIG
ID106	MVYDYFISQQLLFSFLLSTIPT
ID107	MLFLCSCSLSLNQL
ID108	MFFLMVLLFRSNKWT
ID109	MLPLQGLCTCYFLHLEFLSHVTTSLASSS
ID110	MYFYGLTFHFFLLLNTILLFG
ID111	MRWNLFFFCILRNQTKLWASQGSLQDAQS
ID112	MFIAALFTMAKTWN
ID113	MPGXKHFLRVFRXSAXRSVGYXXKPGTSRASLWVXLPXXXVIAS
ID114	MRLESPDENFAVVQEHAIHHIDGPLRRFLLLEVHEPVALGPLFVTGHFA
ID115	MAGSPDREVLLPTVLRGSYC
ID116	MHVSMLEGFDENLDVQGELILQDAFQVWDPKSLIRKGRERHLFLFEISLVFS
ID117	MNVGTAHSEVNXNTRVMKXRGIWLSYVLAIGLLHIVLLS
ID118	MLSFVXAIXXYIPTNS
ID119	MDEYSWWCHVLEVVKGQMFTFINITLWLGSLC
ID120	MRRKGQGHLAFIFLIQIWKTCLS
ID121	MFLISGHVHLIYNILFLAVSSFSMP
ID122	MTPRILSEVQFSAFCPYWTIARILERVGSACFRLELCAAIVGYFVLDVRTFLFIVVCVIC VTLN
ID123	MCSLLSGWGQLLRC
ID124	MLFSFCFPVHFWNPSSLFPPSSVSLIPFNFSASGLCA
ID125	MTWILRILFVIGSXL
ID126	MSSTYCGNSSAKMSVNEVSAFSLSLEQKTGFAFVGILCIFLGLLIIRC
ID127	MTDIWLTMLSMIVGATCY
ID128	MXXCWIYAFISLGYILG
ID129	MFIRTLKTTVLPFMRTAPQLALSWVPPXCRV
ID130	MRTGAEMRTNSSVLIFCLLPYIYH
ID131	MIVIPSWLENEGLELGFSHRTFA
ID132	MLKKEIAHHSPSLVSCPVCTTKYRTLRLLRVISVFLSFLPSYP
ID133	MTXPSRAQTVDXGIAKHCAYSLPGVALTLG
ID134	MPFRLLIPLGLLCALLPQHHG
ID135	MXLVLVFLCSLLAPMVLA
ID136	MALRRPPRLRLCARLPDFFLLLLFRGCLIG
ID137	MGGNGSTCKPDTERQGTLSTAAPTTSPAPCLSNHHNKKHLILAFCAGVLLTLLLIAFIFL

SEQ. ID NO.	SIGNAL PEPTIDE
ID138	MGTADSDEMAPEAPQHTHIDVHIHQESALAKLLLTCCSA
ID139	MSLLSFLFARVNLG
ID140	MARSPLRRRGRPTWSLSTPRPGSPTSSSRSWWCCPARLTLTSG
ID141	MKIITTTLLLACHLQLEVG
ID142	MLMPVVGRGNGIPQTVSEWLRLLPFLGVLALLGYLAV
ID143	MDRLGSFSNDPSDKPPCRGCSSYLMEPYIKCAECGPPPFFLCLQCFTRG
ID144	MSDVNVSALPIKKNSGHIYNKNISQKDCDCLHVVEPMPVRGPDVEAYCLRCECKYEERSS VTIKVTIIIYLSILGLLLLYMVYLTLV
ID145	MAXCRRCRSQRRSHCCQDRRLRRPRLTLWRHHTALSLSLSMAPPNP
ID146	MTRLGGKGGQQFPPGQKIISKDILALTALSVXRKXS
ID147	MKGWGWLALLLGALLGTAWA
ID148	MHRLLCLLLLFGGGDP
ID149	MLWRQLIYWQLLALFFLPFCLC
ID150	MRLLLLLLVAASAMVRS
ID151	MSSGXELLWPGAALLVLLGVAASLC
ID152	MTKEIFFFTVELVCENKELCSSPRWRNAIQKSNFSKVTSFFMSCHHFKGLAPLPHVYTQG NCRPISCLGLTLMPFASS
ID153	MTTDIGCLYFRALCLPRGAWG
ID154	MVPSLVIPDLTCLFLFLNLRWS
ID155	MALRLLKLAATSASA
ID156	MWGNKFGVLLFLYSVLLTKG
ID157	MYTFRKLSPYLNKIVFVCSSVLGQSWG
ID158	MESRVLLRTFCLIFGLGAVWG
ID159	MLVLKKHSVNIAAQTCFKFNFIFRILIFLGFFLGLFH
ID160	MDVKCPGCYKITTVFSHAQTVVLCVGCSTVLC
ID161	MCIILSVLHALPAGIA
ID162	MLVVEASSSVRLASSEVTSWSILVTPSASTPIISLSAGPLRTPSHSKTWLLLGALEPASE
ID163	MYSFPTTVVEEILSLSLQLIAFPTVSC
ID164	MLMLLPLRSLLALVRE
ID165	MVPLVAVVSGPRAQLFACLLRLGTQ
ID166	MDNRFATAFVIACVLSLIST
ID167	MPEYCGNEVTPTEAAQAPEVTYEAEEGSLWTLLLTSLDGHLL
ID168	MNRVLCAPAAGAVRA
ID169	MAFTFAAFCYMLALLLTAALIFF
ID170	MXXXXEXLLAFHHDCEA
ID171	MGPYNVAVPSDVSHARFYFLFHRPLRLLNLLILIEG
ID172	MMNFRQRMGWIGVGLYLLASAAA
ID173	MLFASGGFXVKLYDIEQQQIRNALENIRKEMKLLEQAGSLKGSLSVEEQLSLISGCPNIQ EAVEGAMHIQECVPEDLELKKKIFAQLDSIIDESSDLKRFXFLSHAFXVVCWLGPCEA
ID174	MQCFLGGLGLCSLPLSPSAVCP
ID175	MSSFLLSFSQSLS
ID176	MLTASLAFQLVDG
ID177	MYXRRELSILCILSAFNFLVCLSLG
ID178	MGLSAMDTSIVFGVSWVMLVYS
ID179	MYFWRDVAVSLDTLWALPRQQPGLGNNRVLGLLSGTNKDYKGQKLAEQMFQGIILFSAIV GFIYG
ID180	MHWGKRWXLXXGGLLICXLXIGTATP
ID181	MAXRYNRLTVLAGAXLALGLXTCLSVLFG
ID182	MGFTGFFTATCFISKVFMTCILCRPPISS
ID183	MIMYLFVICVIFEIIRNYAFSILIVLLPVLFFSLK
ID184	MSTVGLXHFPXPLTRICPAPWGLRLWEKLTLLSPGIA
ID185	MLALAXHLSTVES
ID186	MLLSIGMLMLXATQVYTILTVQLFAFLNLLPVEA

SEQ. ID NO.	SIGNAL PEPTIDE
ID187	MSTVGLFHFPTPLTRICPAPWGLRLWEKLTLLSPGIA
ID188	MELTIFILRLAIYILTFPLYLLNFLGLWSWICK
ID189	MSLLHGNKMCVTIRPTGQPLNGDLLLLYLCCMINIHH
ID190	MSFNISYFIAFPNLSQA
ID191	MKLKXNVLTIILLPVHLLIT
ID192	MAALVTVLFTGVRR
ID193	MASVGECPAPVPVKDKKLLEVKLGELPSWILMRDFSPSGIFG
ID194	MLALLVLVTVALASA
ID195	MRIISRQIVLLFSGFWGLAMG

Minimum signal peptide score	false positive rate	false negative rate	proba(0.1)	proba(0.2)
3.5	0.121	0.036	0.467	0.664
4	0.096	0.06	0.519	0.708
4.5	0.078	0.079	0.565	0.745
5	0.062	0.098	0.615	0.782
5.5	0.05	0.127	0.659	0.813
6	0.04	0.163	0.694	0.836
6.5	0.033	0.202	0.725	0.855
7	0.025	0.248	0.763	0.878
7.5	0.021	0.304	0.78	0.889
8	0.015	0.368	0.816	0.909
8.5	0.012	0.418	0.836	0.92
9	0.009	0.512	0.856	0.93
9.5	0.007	0.581	0.863	0.934
10	0.006	0.679	0.835	0.919

TABLE IV

Minimum signal peptide score	All ESTs	New ESTs	ESTs matching public EST closer than 40 bp from beginning	ESTs extending known mRNA more than 40 bp	ESTs extending public EST more than 40 bp
3.5	2674	947	599	23	150
4	2278	784	499	23	126
4.5	1943	647	425	22	112
5	1657	523	353	21	96
5.5	1417	419	307	19	80
6	1190	340	238	18	68
6.5	1035	280	186	18	60
7	893	219	161	15	48
7.5	753	173	132	12	36
8	636	133	101	11	29
8.5	543	104	83	8	26
9	456	81	63	6	24
9.5	364	57	48	6	18
10	303	47	35	6	15

TABLE V

Tissue	All ESTs	New ESTs	ESTs matching public EST closer than 40 bp from beginning	ESTs extending known mRNA more than 40 bp	ESTs extending public EST more than 40 bp
Brain	329	131	75	3	24
Cancerous prostate	134	40	37	1	6
Cerebellum	17	9	1	0	6
Colon	21	11	4	0	0
Dystrophic muscle	41	18	8	0	1
Fetal brain	70	37	16	0	1
Fetal kidney	227	116	46	1	19
Fetal liver	13	7	2	0	0
Heart	30	15	7	0	1
Hypertrophic prostate	86	23	22	2	2
Kidney	10	7	3	0	0
Large intestine	21	8	4	0	1
Liver	23	9	6	0	0
Lung	24	12	4	0	1
Lung (cells)	57	38	6	0	4
Lymph ganglia	163	60	23	2	12
Lymphocytes	23	6	4	0	2
Muscle	33	16	6	0	4
Normal prostate	181	61	45	7	11
Ovary	90	57	12	1	2
Pancreas	48	11	6	0	1
Placenta	24	5	1	0	0
Prostate	34	16	4	0	2
Spleen	56	28	10	0	1
Substantia nigra	108	47	27	1	6
Surrenals	15	3	3	1	0
Testis	131	68	25	1	8
Thyroid	17	8	2	0	2
Umbilical cord	55	17	12	1	3
Uterus	28	15	3	0	2
Non tissue-specific	568	48	177	2	28
Total	2677	947	601	23	150

TABLE VI

Description of Transcription Factor Binding Sites present on promoters isolated from SignalTag sequences

Promoter sequence P13H2 (546 bp):

Matrix	Position	Orientation	Score	Length	Sequence
CMYB_01	-502	+	0.983	9	TGTCAGTTG
MYOD_Q6	-501	-	0.961	10	CCCAACTGAC
S8_01	-444	-	0.960	11	AATAGAATTAG
S8_01	-425	+	0.966	11	AACTAAATTAG
DELTAEF1_01	-390	-	0.960	11	GCACACCTCAG
GATA_C	-364	-	0.964	11	AGATAAATCCA
CMYB_01	-349	+	0.958	9	CTTCAGTTG
GATA1_02	-343	+	0.959	14	TTGTAGATAGGACA
GATA_C	-339	+	0.953	11	AGATAGGACAT
TAL1ALPHAE47_01	-235	+	0.973	16	CATAACAGATGGTAAG
TAL1BETAE47_01	-235	+	0.983	16	CATAACAGATGGTAAG
TAL1BETAITF2_01	-235	+	0.978	16	CATAACAGATGGTAAG
MYOD_Q6	-232	-	0.954	10	ACCATCTGTT
GATA1_04	-217	-	0.953	13	TCAAGATAAAGTA
IK1_01	-126	+	0.963	13	AGTTGGGAATTCC
IK2_01	-126	+	0.985	12	AGTTGGGAATTC
CREL_01	-123	+	0.962	10	TGGGAATTCC
GATA1_02	-96	+	0.950	14	TCAGTGATATGGCA
SRY_02	-41	-	0.951	12	TAAAACAAAACA
E2F_02	-33	+	0.957	8	TTTAGCGC
MZF1_01	-5	-	0.975	8	TGAGGGGA

Promoter sequence P15B4 (861bp) :

Matrix	Position	Orientation	Score	Length	Sequence
NFY_Q6	-748	-	0.956	11	GGACCAATCAT
MZF1_01	-738	+	0.962	8	CCTGGGGA
CMYB_01	-684	+	0.994	9	TGACCGTTG
VMYB_02	-682	-	0.985	9	TCCAACGGT
STAT_01	-673	+	0.968	9	TTCCTGGAA
STAT_01	-673	-	0.951	9	TTCCAGGAA
MZF1_01	-556	-	0.956	8	TTGGGGGA
IK2_01	-451	+	0.965	12	GAATGGGATTTC
MZF1_01	-424	+	0.986	8	AGAGGGGA
SRY_02	-398	-	0.955	12	GAAAACAAAACA
MZF1_01	-216	+	0.960	8	GAAGGGGA
MYOD_Q6	-190	+	0.981	10	AGCATCTGCC
DELTAEF1_01	-176	+	0.958	11	TCCCACCTTCC
S8_01	5	-	0.992	11	GAGGCAATTAT
MZF1_01	16	-	0.986	8	AGAGGGGA

Promoter sequence P29B6 (555 bp) :

Matrix	Position	Orientation	Score	Length	Sequence
ARNT_01	-311	+	0.964	16	GGACTCACGTGCTGCT
NMYC_01	-309	+	0.965	12	ACTCACGTGCTG
USF_01	-309	+	0.985	12	ACTCACGTGCTG
USF_01	-309	-	0.985	12	CAGCACGTGAGT
NMYC_01	-309	-	0.956	12	CAGCACGTGAGT
MYCMAX_02	-309	-	0.972	12	CAGCACGTGAGT
USF_C	-307	+	0.997	8	TCACGTGC
USF_C	-307	-	0.991	8	GCACGTGA
MZF1_01	-292	-	0.968	8	CATGGGGA
ELK1_02	-105	+	0.963	14	CTCTCCGGAAGCCT
CETS1P54_01	-102	+	0.974	10	TCCGGAAGCC
AP1_Q4	-42	-	0.963	11	AGTGACTGAAC
AP1FJ_Q2	-42	-	0.961	11	AGTGACTGAAC
PADS_C	45	+	1.000	9	TGTGGTCTC

TABLE VII

CLAIMS

1. A purified or isolated nucleic acid comprising the sequence of one of SEQ ID NOs: 38-195 or comprising a sequence complementary thereto.

2. The nucleic acid of Claim 1, wherein said nucleic acid is recombinant.

3. A purified or isolated nucleic acid comprising at least 10 consecutive bases of the sequence of one of SEQ ID NOs: 38-195 or one of the sequences complementary thereto.

4. A purified or isolated nucleic acid comprising at least 15 consecutive bases of one of the sequences of SEQ ID NOs: 38-195 or one of the sequences complementary thereto.

5. The nucleic acid of Claim 4, wherein said nucleic acid is recombinant.

6. A purified or isolated nucleic acid of at least 15 bases capable of hybridizing under stringent conditions to the sequence of one of SEQ ID NOs: 38-195 or one of the sequences complementary to the sequences of SEQ ID NOs: 38-195.

7. The nucleic acid of Claim 6, wherein said nucleic acid is recombinant.

8. A purified or isolated nucleic acid encoding a human gene product, said human gene product having a sequence partially encoded by one of the sequences of SEQ ID NO: 38-195.

9. A purified or isolated nucleic acid having the sequence of one of SEQ ID NOs: 38-195 or having a sequence complementary thereto.

10. A purified or isolated nucleic acid comprising the nucleotides of one of SEQ ID NOs: 38-195 which encode a signal peptide.

11. A purified or isolated polypeptides comprising a signal peptide encoded by one of the sequences of SEQ ID NOs: 38-195.

12. A vector encoding a fusion protein comprising a polypeptide and a signal peptide, said vector comprising a first nucleic acid encoding a signal peptide encoded by one of the sequences of SEQ ID NOs: 38-195 operably linked to a second nucleic acid encoding a polypeptide.

13. A method of directing the extracellular secretion of a polypeptide or the insertion of a polypetide into the membrane comprising the steps of:

obtaining a vector according to Claim 12; and

introducing said vector into a host cell such that said fusion protein is secreted into the extracellular environment of said host cell or inserted into the membrane of said host cell.

14. A method of importing a polypeptide into a cell comprising contacting said cell with a fusion protein comprising a signal peptide encoded by one of the sequences of SEQ ID NOs: 38-195 operably linked to said polypeptide.

15. A method of making a cDNA encoding a human secretory protein that is partially encoded by one of SEQ ID NOs 38-195, comprising the steps of:

obtaining a cDNA comprising one of the sequences of SEQ ID NOs: 38-195;

contacting said cDNA with a detectable probe comprising at least 15 consecutive nucleotides of said sequence of SEQ ID NO: 38-195 or a sequence complementary thereto under conditions which permit said probe to hybridize to said cDNA;

identifying a cDNA which hybridizes to said detectable probe; and

isolating said cDNA which hybridizes to said probe.

16. An isolated or purified cDNA encoding a human secretory protein, said human secretory protein comprising the protein encoded by one of SEQ ID NOs 38-195 or a fragment thereof of at least 10 amino acids, said cDNA being obtainable by the method of Claim 15.

17. The cDNA of Claim 16 wherein said cDNA comprises the full protein coding sequence partially included in one of the sequences of SEQ ID NOs: 38-195.

18. A method of making a cDNA comprising one of the sequences of SEQ ID NOs: 38-195, comprising the steps of:

contacting a collection of mRNA molecules from human cells with a first primer capable of hybridizing to the polyA tail of said mRNA;

hybridizing said first primer to said polyA tail;

reverse transcribing said mRNA to make a first cDNA strand;

making a second cDNA strand complementary to said first cDNA strand using at least one primer comprising at least 15 nucleotides of one of the sequences of SEQ ID NOs 38-195; and

isolating the resulting cDNA comprising said first cDNA strand and said second cDNA strand.

19. An isolated or purified cDNA encoding a human secretory protein, said human secretory protein comprising the protein encoded by one of SEQ ID NOs 38-195 or a fragment thereof of at least 10 amino acids, said cDNA being obtainable by the method of Claim 18.

20. The cDNA of Claim 19 wherein said cDNA comprises the full protein coding sequence partially included in one of the sequences of SEQ ID NOs: 38-195.

21. The method of Claim 18, wherein the second cDNA strand is made by:

contacting said first cDNA strand with a first pair of primers, said first pair of primers comprising a second primer comprising at least 15 consecutive nucleotides of one of the sequences of SEQ ID NOs 38-195 and a third primer having a sequence therein which is included within the sequence of said first primer;

performing a first polymerase chain reaction with said first pair of nested primers to generate a first PCR product;

contacting said first PCR product with a second pair of primers, said second pair of primers comprising a fourth primer, said fourth primer comprising at least 15 consecutive nucleotides of said sequence of one of SEQ ID NO:s 38-195 , and a fifth primer, said fourth and fifth primers being capable of hybridizing to sequences within said first PCR product; and

performing a second polymerase chain reaction, thereby generating a second PCR product.

22. An isolated or purified cDNA encoding a human secretory protein, said human secretory protein comprising the protein encoded by one of SEQ ID NOs 38-195, or a fragment thereof of at least 10 amino acids, said cDNA being obtainable by the method of Claim 21.

23. The cDNA of Claim 22 wherein said cDNA comprises the full protein coding sequence partially included in one of the sequences of SEQ ID NOs: 38-195.

24. The method of Claim 18 wherein the second cDNA strand is made by:

contacting said first cDNA strand with a second primer comprising at least 15 consecutive nucleotides of the sequences of SEQ ID NOs: 38-195;

hybridizing said second primer to said first strand cDNA; and

extending said hybridized second primer to generate said second cDNA strand.

25. An isolated or purified cDNA encoding a human secretory protein, said human secretory protein comprising the protein partially encoded by one of SEQ ID NOs 38-195 or comprising a fragment thereof of at least 10 amino acids, said cDNA being obtainable by the method of Claim 24.

26. The cDNA of Claim 25, wherein said cDNA comprises the full protein coding sequence partially included in of one of the sequences of SEQ ID NOs: 38-195.

27. A method of making a protein comprising one of the sequences of SEQ ID NO: 196-353, comprising the steps of:

obtaining a cDNA encoding the full protein sequence partially included in one of the sequences of sequence of SEQ ID NO: 38-195;

inserting said cDNA in an expression vector such that said cDNA is operably linked to a promoter;

introducing said expression vector into a host cell whereby said host cell produces the protein encoded by said cDNA; and

isolating said protein.

28. An isolated protein obtainable by the method of Claim 27.

29. A method of obtaining a promoter DNA comprising the steps of:

obtaining DNAs located upstream of the nucleic acids of SEQ ID NO: 38-195 or the sequences complementary thereto;

screening said upstream DNAs to identify a promoter capable of directing transcription initiation; and

isolating said DNA comprising said identified promoter.

30. The method of Claim 29, wherein said obtaining step comprises chromosome walking from said nucleic acids of SEQ ID NO: 38-195 or sequences complementary thereto.

31. The method of Claim 30, wherein said screening step comprises inserting said upstream sequences into a promoter reporter vector.

32. The method of Claim 30, wherein said screening step comprises identifying motifs in said upstream DNAs which are transcription factor binding sites or transcription start sites.

33. An isolated promoter obtainable by the method of Claim 32.

34. An isolated or purified protein comprising one of the sequences of SEQ ID NO: 196-353.

35. In an array of discrete ESTs or fragments thereof of at least 15 nucleotides in length, the improvement comprising inclusion in said array of at least one of the sequences of SEQ ID NOs: 38-195, or one of the sequences complementary to the sequences of SEQ ID NOs: 38-195, or a fragment thereof of at least 15 consecutive nucleotides.

36. The array of Claim 35 including therein at least two of the sequences of SEQ ID NOs: 38-195, the sequences complementary to the sequences of SEQ ID NOs: 38-195, or fragments thereof of at least 15 consecutive nucleotides.

37. The array of Claim 35 including therein at least five of the sequences of SEQ ID NOs: 38-195, the sequences complementary to the sequences of SEQ ID NOs: 38-195, or fragments thereof of at least 15 consecutive nucleotides.

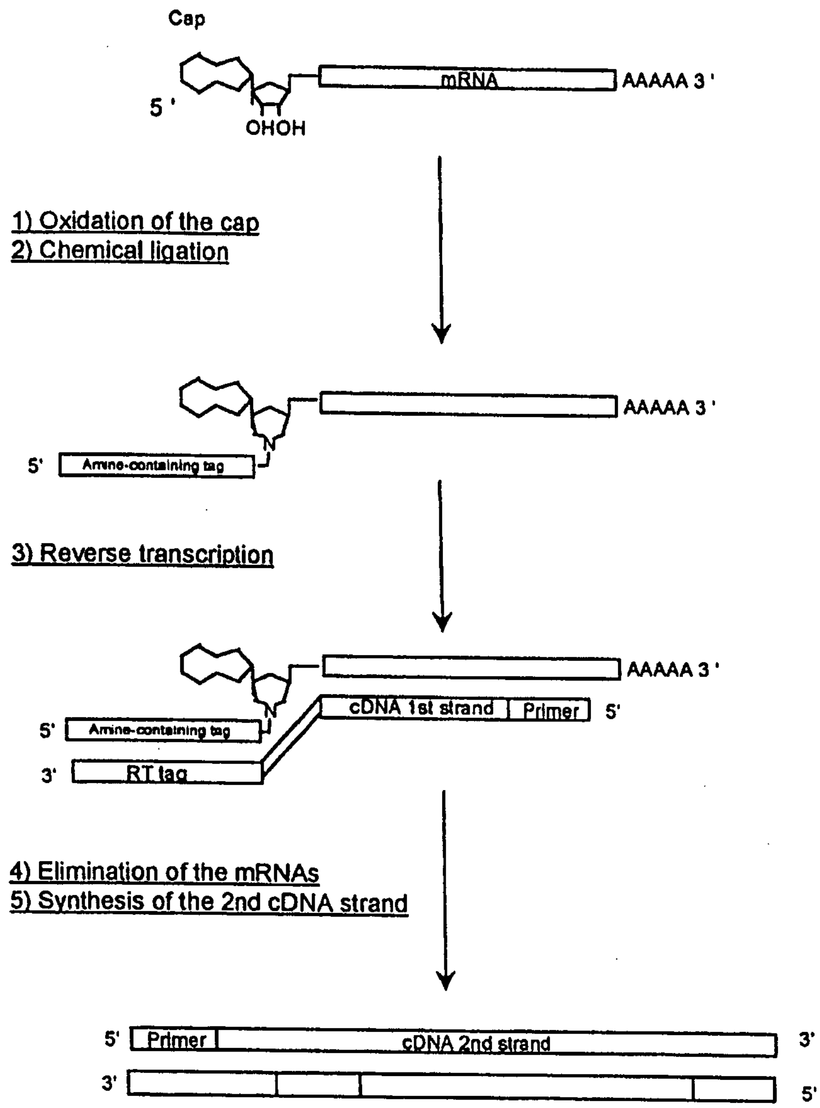

Cap
5'
OHOH
mRNA
AAAAA 3'
1) Oxidation of the cap
2) Chemical ligation
AAAAA 3'
5'
Amine-containing tag
3) Reverse transcription
AAAAA 3'
cDNA 1st strand
Primer
5'
5'
Amine-containing tag
3'
RT tag
4) Elimination of the mRNAs
5) Synthesis of the 2nd cDNA strand
5'
Primer
cDNA 2nd strand
3'
3'
5'

Figure 1

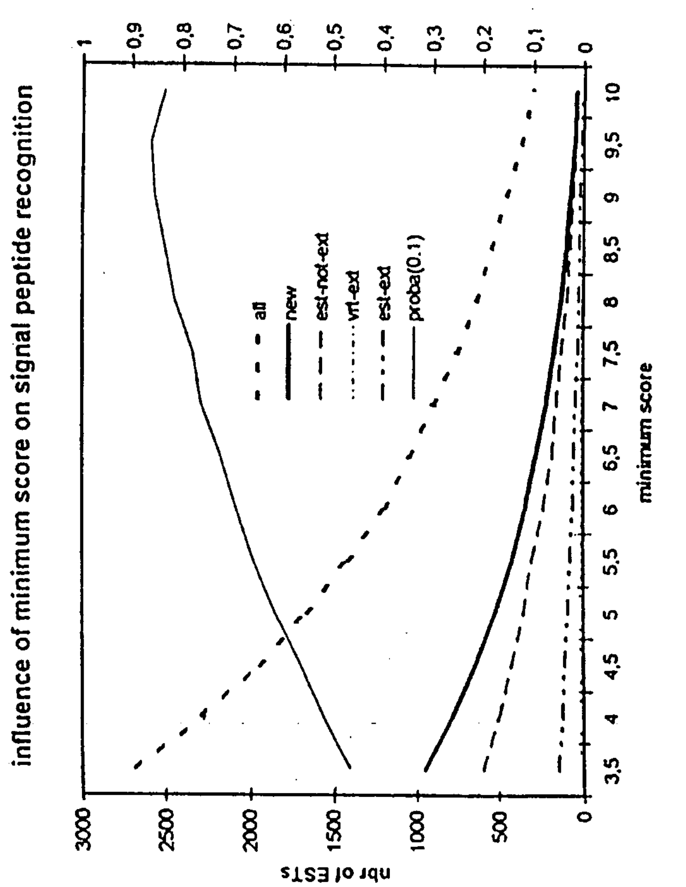

influence of minimum score on signal peptide recognition
nbr of ESTs
3000
2500
2000
1500
1000
500
0
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
3,5
4
4,5
5
5,5
6
6,5
7
7,5
8
8,5
9
9,5
10
minimum score
all
new
est-not-ext
vrt-ext
est-ext
proba(0.1)

Figure 2

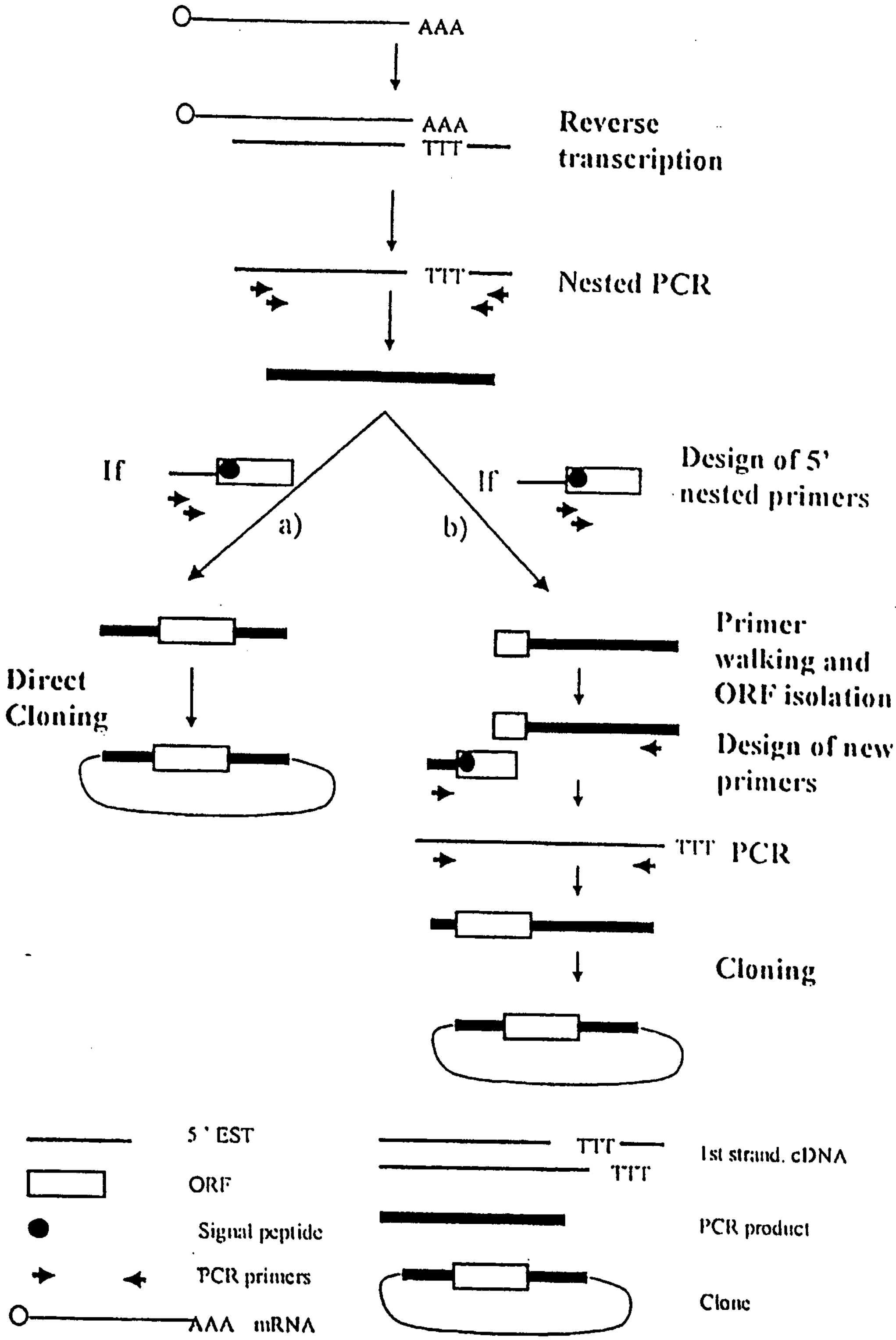

AAA
AAA
TTT
Reverse transcription
TTT
Nested PCR
If
If
a)
b)
Design of 5' nested primers
Direct Cloning
Primer walking and ORF isolation
Design of new primers
TTT
PCR
Cloning
5 ' EST
ORF
Signal peptide
PCR primers
AAA mRNA
TTT
TTT
1st strand. cDNA
PCR product
Clone

Figure 3

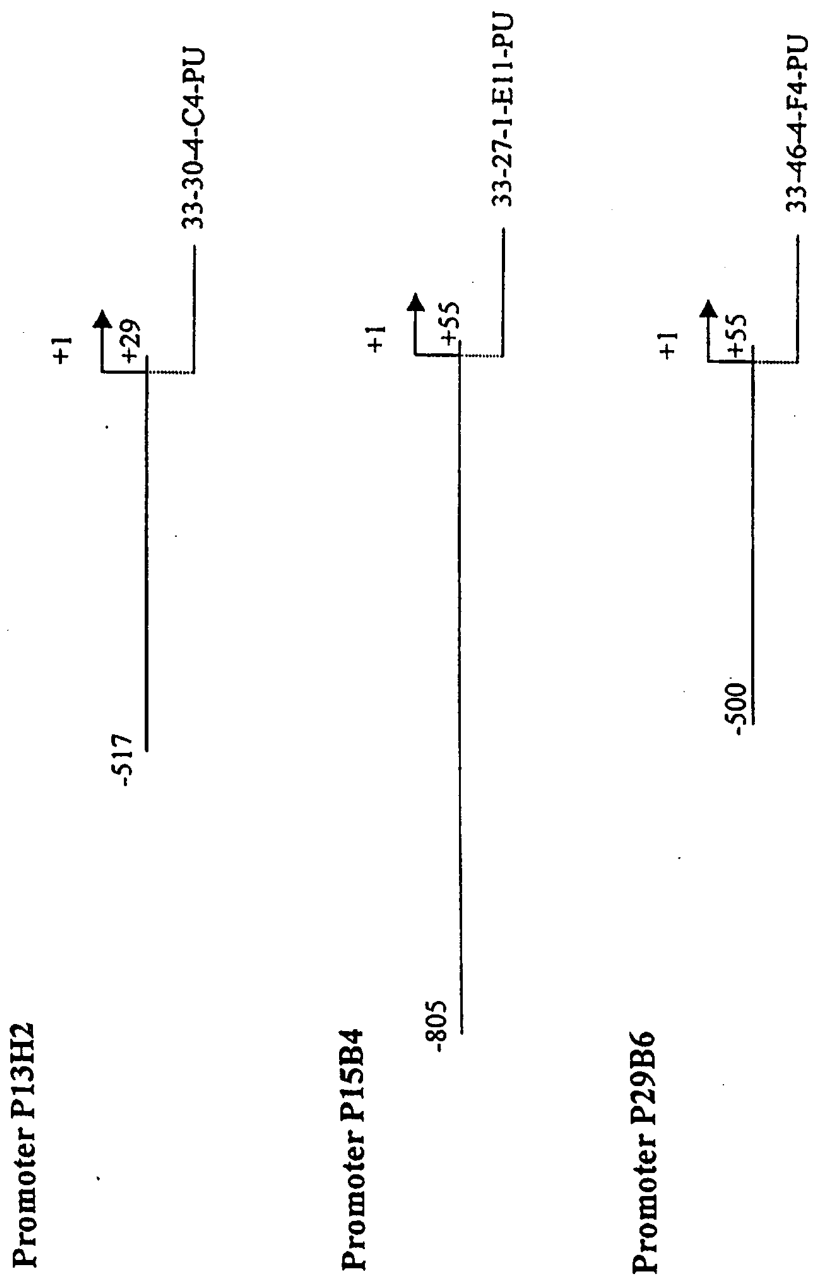

Promoter P13H2
-517
+1
+29
33-30-4-C4-PU
Promoter P15B4
-805
+1
+55
33-27-1-E11-PU
Promoter P29B6
-500
+1
+55
33-46-4-F4-PU

Figure 4

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME : GENSET SA
(B) STREET :24, RUE ROYALE
(C) CITY: PARIS
(E) COUNTRY : FRANCE
(F) POSTAL CODE (ZIP) : 75008

(ii) TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS IDENTIFIED FROM BRAIN TISSUE

(iii) NUMBER OF SEQUENCES: 353

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy Disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: Win95
(D) SOFTWARE: Word

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 47 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(ix) FEATURE:
(A) NAME/KEY: Cap
(B) LOCATION: 1
(D) OTHER INFORMATION: m7Gppp added to 1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GGCAUCCUAC UCCCAUCCAA UUCCACCCUA ACUCCUCCCA UCUCCAC 47

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 46 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GCAUCCUACU CCCAUCCAAU UCCACCCUAA CUCCUCCCAU CUCCAC 46

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

ATCAAGAATT CGCACGAGAC CATTA 25

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

TAATGGTCTC GTGCGAATTC TTGAT 25

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

CCGACAAGAC CAACGTCAAG GCCGC 25

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID

(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

TCACCAGCAG GCAGTGGCTT AGGAG 25

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

AGTGATTCCT GCTACTTTGG ATGGC 25

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GCTTGGTCTT GTTCTGGAGT TTAGA 25

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

TCCAGAATGG GAGACAAGCC AATTT 25

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

AGGGAGGAGG AAACAGCGTG AGTCC 25

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

ATGGGAAAGG AAAAGACTCA TATCA 25

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

AGCAGCAACA ATCAGGACAG CACAG 25

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

ATCAAGAATT CGCACGAGAC CATTA 25

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 67 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

ATCGTTGAGA CTCGTACCAG CAGAGTCACG AGAGAGACTA CACGGTACTG GTTTTTTTTT 60

TTTTTVN 67

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

CCAGCAGAGT CACGAGAGAG ACTACACGG 29

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

CACGAGAGAG ACTACACGGT ACTGG 25

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 526 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lymph ganglia

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(261..376)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 166..281
id N70479
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(380..486)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 54..160
id N70479
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(110..145)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 403..438
id N70479
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(196..229)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 315..348
id N70479
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 90..140
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.2
seq LLLITAILAVAVG/FP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

```
AATATRARAC AGCTACAATA TTCCAGGGCC ARTCACTTGC CATTTCTCAT AACAGCGTCA     60

GAGAGAAAGA ACTGACTGAR ACGTTTGAG ATG AAG AAA GTT CTC CTC CTG ATC      113
                                Met Lys Lys Val Leu Leu Leu Ile
                                        -15                 -10

ACA GCC ATC TTG GCA GTG GCT GTW GGT TTC CCA GTC TCT CAA GAC CAG      161
Thr Ala Ile Leu Ala Val Ala Val Gly Phe Pro Val Ser Gln Asp Gln
                -5                  1                   5

GAA CGA GAA AAA AGA AGT ATC AGT GAC AGC GAT GAA TTA GCT TCA GGR      209
Glu Arg Glu Lys Arg Ser Ile Ser Asp Ser Asp Glu Leu Ala Ser Gly
        10                  15                  20

WTT TTT GTG TTC CCT TAC CCA TAT CCA TTT CGC CCA CTT CCA CCA ATT      257
Xaa Phe Val Phe Pro Tyr Pro Tyr Pro Phe Arg Pro Leu Pro Pro Ile
    25                  30                  35

CCA TTT CCA AGA TTT CCA TGG TTT AGA CGT AAN TTT CCT ATT CCA ATA      305
Pro Phe Pro Arg Phe Pro Trp Phe Arg Arg Xaa Phe Pro Ile Pro Ile
40                  45                  50                  55

CCT GAA TCT GCC CCT ACA ACT CCC CTT CCT AGC GAA AAG TAAACAARAA       354
Pro Glu Ser Ala Pro Thr Thr Pro Leu Pro Ser Glu Lys
                60                  65

GGAAAAGTCA CRATAAACCT GGTCACCTGA AATTGAAATT GAGCCACTTC CTTGAARAAT    414

CAAAATTCCT GTTAATAAAA RAAAAACAAA TGTAATTGAA ATAGCACACA GCATTCTCTA    474

GTCAATATCT TTAGTGATCT TCTTTAATAA ACATGAAAGC AAAAAAAAAA AA            526
```

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 1..17
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.2
seq LLLITAILAVAVG/FP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

```
Met Lys Lys Val Leu Leu Leu Ile Thr Ala Ile Leu Ala Val Ala Val
1               5                   10                  15

Gly
```

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 822 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 260..464
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 153..357
id H57434
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 118..184
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 98..164
id H57434
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 56..113
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 35..92
id H57434
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 454..485
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 348..379
id H57434
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 118..545
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..428
id N27248
est

```
(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 65..369
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 98
                             region 41..345
                             id H94779
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 61..399
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 99
                             region 6..344
                             id H09880
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 408..458
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 92
                             region 355..405
                             id H09880
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 60..399
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 97
                             region 56..395
                             id H29351
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 393..432
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 90
                             region 391..430
                             id H29351
                             est

(ix) FEATURE:
     (A) NAME/KEY: sig_peptide
     (B) LOCATION: 346..408
     (C) IDENTIFICATION METHOD: Von Heijne matrix
     (D) OTHER INFORMATION:  score 5.5
                             seq SFLPSALVIWTSA/AF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

ACTCCTTTTA GCATAGGGGC TTCGGCGCCA GCGGCCAGCG CTAGTCGGTC TGGTAAGTGC     60

CTGATGCCGA GTTCCGTCTC TCGCGTCTTT TCCTGGTCCC AGGCAAAGCG GASGNAGATC    120

CTCAAACGGC CTAGTGCTTC GCGCTTCCGG AGAAAATCAG CGGTCTAATT AATTCCTCTG    180
```

```
GTTTGTTGAA GCAGTTACCA AGAATCTTCA ACCCTTTCCC ACAAAAGCTA ATTGAGTACA    240

CGTTCCTGTT GAGTACACGT TCCTGTTGAT TTACAAAAGG TGCAGGTATG AGCAGGTCTG    300

AAGACTAACA TTTTGTGAAG TTGTAAAACA GAAAACCTGT TAGAA ATG TGG TGG TTT    357
                                                  Met Trp Trp Phe
                                                      -20

CAG CAA GGC CTC AGT TTC CTT CCT TCA GCC CTT GTA ATT TGG ACA TCT      405
Gln Gln Gly Leu Ser Phe Leu Pro Ser Ala Leu Val Ile Trp Thr Ser
        -15                 -10                 -5

GCT GCT TTC ATA TTT TCA TAC ATT ACT GCA GTA ACA CTC CAC CAT ATA      453
Ala Ala Phe Ile Phe Ser Tyr Ile Thr Ala Val Thr Leu His His Ile
    1               5                   10                  15

GAC CCG GCT TTA CCT TAT ATC AGT GAC ACT GGT ACA GTA GCT CCA RAA      501
Asp Pro Ala Leu Pro Tyr Ile Ser Asp Thr Gly Thr Val Ala Pro Xaa
                20                  25                  30

AAA TGC TTA TTT GGG GCA ATG CTA AAT ATT GCG GCA GTT TTA TGT CAA      549
Lys Cys Leu Phe Gly Ala Met Leu Asn Ile Ala Ala Val Leu Cys Gln
            35                  40                  45

AAA TAGAAATCAG GAARATAATT CAACTTAAAG AAKTTCATTT CATGACCAAA           602
Lys

CTCTTCARAA ACATGTCTTT ACAAGCATAT CTCTTGTATT GCTTTCTACA CTGTTGAATT    662

GTCTGGCAAT ATTTCTGCAG TGGAAAATTT GATTTARMTA GTTCTTGACT GATAAATATG    722

GTAAGGTGGG CTTTTCCCCC TGTGTAATTG GCTACTATGT CTTACTGAGC CAAGTTGTAW    782

TTTGAAATAA AATGATATGA GAGTGACACA AAAAAAAAAA                          822
```

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 1..21
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.5
seq SFLPSALVIWTSA/AF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

```
Met Trp Trp Phe Gln Gln Gly Leu Ser Phe Leu Pro Ser Ala Leu Val
  1               5                  10                  15

Ile Trp Thr Ser Ala
            20
```

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 405 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Testis

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(103..398)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 1..296
id AA442893
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 185..295
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.9
seq LSYASSALSPCLT/AP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

```
ATCACCTTCT TCTCCATCCT TSTCTGGGCC AGTCCCCARC CCAGTCCCTC TCCTGACCTG     60

CCCAGCCCAA GTCAGCCTTC AGCACGCGCT TTTCTGCACA CAGATATTCC AGGCCTACCT    120

GGCATTCCAG GACCTCCGMA ATGATGCTCC AGTCCCTTAC AAGCGCTTCC TGGATGAGGG    180

TGGC ATG GTG CTG ACC ACC CTC CCC TTG CCC TCT GCC AAC AGC CCT GTG     229
     Met Val Leu Thr Thr Leu Pro Leu Pro Ser Ala Asn Ser Pro Val
             -35                 -30                 -25

AAC ATG CCC ACC ACT GGC CCC AAC AGC CTG AGT TAT GCT AGC TCT GCC      277
Asn Met Pro Thr Thr Gly Pro Asn Ser Leu Ser Tyr Ala Ser Ser Ala
        -20                 -15                 -10

CTG TCC CCC TGT CTG ACC GCT CCA AAK TCC CCC CGG CTT GCT ATG ATG      325
Leu Ser Pro Cys Leu Thr Ala Pro Xaa Ser Pro Arg Leu Ala Met Met
    -5                  1               5                  10

CCT GAC AAC TAAATATCCT TATCCAAATC AATAAARWRA RAATCCTCCC TCCARAAGGG   384
Pro Asp Asn
```

```
TTTCTAAAAA CAAAAAAAAA A                                                405

(2) INFORMATION FOR SEQ ID NO: 22:

      (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 37 amino acids
            (B) TYPE: AMINO ACID
            (D) TOPOLOGY: LINEAR

      (ii) MOLECULE TYPE: PROTEIN

      (vi) ORIGINAL SOURCE:
            (A) ORGANISM: Homo Sapiens

      (ix) FEATURE:
            (A) NAME/KEY: sig_peptide
            (B) LOCATION: 1..37
            (C) IDENTIFICATION METHOD: Von Heijne matrix
            (D) OTHER INFORMATION:  score 5.9
                                    seq LSYASSALSPCLT/AP

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

Met Val Leu Thr Thr Leu Pro Leu Pro Ser Ala Asn Ser Pro Val Asn
  1               5                  10                  15

Met Pro Thr Thr Gly Pro Asn Ser Leu Ser Tyr Ala Ser Ser Ala Leu
             20                  25                  30

Ser Pro Cys Leu Thr
         35


(2) INFORMATION FOR SEQ ID NO: 23:

      (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 496 base pairs
            (B) TYPE: NUCLEIC ACID
            (C) STRANDEDNESS: DOUBLE
            (D) TOPOLOGY: LINEAR

      (ii) MOLECULE TYPE: CDNA

      (vi) ORIGINAL SOURCE:
            (A) ORGANISM: Homo Sapiens
            (F) TISSUE TYPE: Cancerous prostate

      (ix) FEATURE:
            (A) NAME/KEY: other
            (B) LOCATION: 149..331
            (C) IDENTIFICATION METHOD: blastn
            (D) OTHER INFORMATION:  identity 98
                                    region 1..183
                                    id AA397994
                                    est
```

```
(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 328..485
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 96
                              region 179..336
                              id AA397994
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: complement(182..496)
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 97
                              region 14..328
                              id AA399680
                              est

(ix) FEATURE:
      (A) NAME/KEY: sig_peptide
      (B) LOCATION: 196..240
      (C) IDENTIFICATION METHOD: Von Heijne matrix
      (D) OTHER INFORMATION:  score 5.5
                              seq ILSTVTALTFAXA/LD

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

AAAAAATTGG TCCCAGTTTT CACCCTGCCG CAGGGCTGGC TGGGGAGGGC AGCGGTTTAG      60

ATTAGCCGTG GCCTAGGCCG TTTAACGGGG TGACACGAGC NTGCAGGGCC GAGTCCAAGG     120

CCCGGAGATA GGACCAACCG TCAGGAATGC GAGGAATGTT TTTCTTCGGA CTCTATCGAG     180

GCACACAGAC AGACC ATG GGG ATT CTG TCT ACA GTG ACA GCC TTA ACA TTT      231
                 Met Gly Ile Leu Ser Thr Val Thr Ala Leu Thr Phe
                 -15                 -10                 -5

GCC ARA GCC CTG GAC GGC TGC AGA AAT GGC ATT GCC CAC CCT GCA AGT       279
Ala Xaa Ala Leu Asp Gly Cys Arg Asn Gly Ile Ala His Pro Ala Ser
            1               5                  10

GAG AAG CAC AGA CTC GAG AAA TGT AGG GAA CTC GAG ASC ASC CAC TCG       327
Glu Lys His Arg Leu Glu Lys Cys Arg Glu Leu Glu Xaa Xaa His Ser
    15                  20                  25

GCC CCA GGA TCA ACC CAS CAC CGA AGA AAA ACA ACC AGA AGA AAT TAT       375
Ala Pro Gly Ser Thr Xaa His Arg Arg Lys Thr Thr Arg Arg Asn Tyr
30                  35                  40                  45

TCT TCA GCC TGAAATGAAK CCGGGATCAA ATGGTTGCTG ATCARAGCCC ATATTTAAAT    434
Ser Ser Ala

TGGAAAAGTC AAATTGASCA TTATTAAATA AAGCTTGTTT AATATGTCTC AAACAAAAAA     494

AA                                                                    496
```

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 1..15
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.5
seq ILSTVTALTFAXA/LD

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

```
Met Gly Ile Leu Ser Thr Val Thr Ala Leu Thr Phe Ala Xaa Ala
  1               5                  10                  15
```

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 623 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Testis

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 49..96
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 10.1
seq LVLTLCTLPLAVA/SA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

```
AAAGATCCCT GCAGCCCGGC AGGAGAGAAG GCTGAGCCTT CTGGCGTC ATG GAG AGG      57
                                                     Met Glu Arg
                                                         -15

CTC GTC CTA ACC CTG TGC ACC CTC CCG CTG GCT GTG GCG TCT GCT GGC      105
Leu Val Leu Thr Leu Cys Thr Leu Pro Leu Ala Val Ala Ser Ala Gly
            -10                 -5                  1

TGC GCC ACG ACG CCA GCT CGC AAC CTG AGC TGC TAC CAG TGC TTC AAG      153
Cys Ala Thr Thr Pro Ala Arg Asn Leu Ser Cys Tyr Gln Cys Phe Lys
```

```
    5                   10                  15

GTC AGC AGC TGG ACG GAG TGC CCG CCC ACC TGG TGC AGC CCG CTG GAC      201
Val Ser Ser Trp Thr Glu Cys Pro Pro Thr Trp Cys Ser Pro Leu Asp
20                  25                  30                  35

CAA GTC TGC ATC TCC AAC GAG GTG GTC GTC TCT TTT AAA TGG AGT GTA      249
Gln Val Cys Ile Ser Asn Glu Val Val Val Ser Phe Lys Trp Ser Val
                40                  45                  50

CGC GTC CTG CTC AGC AAA CGC TGT GCT CCC AGA TGT CCC AAC GAC AAC      297
Arg Val Leu Leu Ser Lys Arg Cys Ala Pro Arg Cys Pro Asn Asp Asn
            55                  60                  65

ATG AAK TTC GAA TGG TCG CCG GCC CCC ATG GTG CAA GGC GTG ATC ACC      345
Met Xaa Phe Glu Trp Ser Pro Ala Pro Met Val Gln Gly Val Ile Thr
        70                  75                  80

AGG CGC TGC TGT TCC TGG GCT CTC TGC AAC AGG GCA CTG ACC CCA CAG      393
Arg Arg Cys Cys Ser Trp Ala Leu Cys Asn Arg Ala Leu Thr Pro Gln
    85                  90                  95

GAG GGG CGC TGG GCC CTG CRA GGG GGG CTC CTG CTC CAG GAC CCT TCG      441
Glu Gly Arg Trp Ala Leu Xaa Gly Gly Leu Leu Leu Gln Asp Pro Ser
100                 105                 110                 115

AGG GGC ARA AAA ACC TGG GTG CGG CCA CAG CTG GGG CTC CCA CTC TGC      489
Arg Gly Xaa Lys Thr Trp Val Arg Pro Gln Leu Gly Leu Pro Leu Cys
                120                 125                 130

CTT CCC AWT TCC AAC CCC CTC TGC CCA RGG GAA ACC CAG GAA GGA          534
Leu Pro Xaa Ser Asn Pro Leu Cys Pro Xaa Glu Thr Gln Glu Gly
            135                 140                 145

TAACACTGTG GGTGCCCCCA CCTGTGCATT GGGACCACRA CTTCACCCTC TTGGARACAA    594

TAAACTCTCA TGCCCCCAAA AAAAAAAAA                                      623
```

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 1..16
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 10.1
seq LVLTLCTLPLAVA/SA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

```
Met Glu Arg Leu Val Leu Thr Leu Cys Thr Leu Pro Leu Ala Val Ala
  1               5                  10                  15
```

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 848 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 32..73
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 10.7
seq LWLLFFLVTAIHA/EL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

```
AACTTTGCCT TGTGTTTTCC ACCCTGAAAG A ATG TTG TGG CTG CTC TTT TTT CTG      55
                                   Met Leu Trp Leu Leu Phe Phe Leu
                                                   -10

GTG ACT GCC ATT CAT GCT GAA CTC TGT CAA CCA GGT GCA GAA AAT GCT        103
Val Thr Ala Ile His Ala Glu Leu Cys Gln Pro Gly Ala Glu Asn Ala
     -5                  1                5                  10

TTT AAA GTG AGA CTT AGT ATC AGA ACA GCT CTG GGA GAT AAA GCA TAT        151
Phe Lys Val Arg Leu Ser Ile Arg Thr Ala Leu Gly Asp Lys Ala Tyr
                15                  20                  25

GCC TGG GAT ACC AAT GAA GAA TAC CTC TTC AAA GCG ATG GTA GCT TTC        199
Ala Trp Asp Thr Asn Glu Glu Tyr Leu Phe Lys Ala Met Val Ala Phe
            30                  35                  40

TCC ATG AGA AAA GTT CCC AAC AGA GAA GCA ACA GAA ATT TCC CAT GTC        247
Ser Met Arg Lys Val Pro Asn Arg Glu Ala Thr Glu Ile Ser His Val
        45                  50                  55

CTA CTT TGC AAT GTA ACC CAG AGG GTA TCA TTC TGG TTT GTG GTT ACA        295
Leu Leu Cys Asn Val Thr Gln Arg Val Ser Phe Trp Phe Val Val Thr
     60                  65                  70

GAC CCT TCA AAA AAT CAC ACC CTT CCT GCT GTT GAG GTG CAA TCA GCC        343
Asp Pro Ser Lys Asn His Thr Leu Pro Ala Val Glu Val Gln Ser Ala
75                  80                  85                  90

ATA AGA ATG AAC AAG AAC CGG ATC AAC AAT GCC TTC TTT CTA AAT GAC        391
Ile Arg Met Asn Lys Asn Arg Ile Asn Asn Ala Phe Phe Leu Asn Asp
```

```
                95                  100                 105

CAA ACT CTG GAA TTT TTA AAA ATC CCT TCC ACA CTT GCA CCA CCC ATG       439
Gln Thr Leu Glu Phe Leu Lys Ile Pro Ser Thr Leu Ala Pro Pro Met
            110                 115                 120

GAC CCA TCT GTG CCC ATC TGG ATT ATT ATA TTT GGT GTG ATA TTT TGC       487
Asp Pro Ser Val Pro Ile Trp Ile Ile Ile Phe Gly Val Ile Phe Cys
        125                 130                 135

ATC ATC ATA GTT GCA ATT GCA CTA CTG ATT TTA TCA GGG ATC TGG CAA       535
Ile Ile Ile Val Ala Ile Ala Leu Leu Ile Leu Ser Gly Ile Trp Gln
    140                 145                 150

CGT ADA ARA AAG AAC AAA GAA CCA TCT GAA GTG GAT GAC GCT GAA RAT       583
Arg Xaa Xaa Lys Asn Lys Glu Pro Ser Glu Val Asp Asp Ala Glu Xaa
155                 160                 165                 170

AAK TGT GAA AAC ATG ATC ACA ATT GAA AAT GGC ATC CCC TCT GAT CCC       631
Xaa Cys Glu Asn Met Ile Thr Ile Glu Asn Gly Ile Pro Ser Asp Pro
                175                 180                 185

CTG GAC ATG AAG GGA GGG CAT ATT AAT GAT GCC TTC ATG ACA GAG GAT       679
Leu Asp Met Lys Gly Gly His Ile Asn Asp Ala Phe Met Thr Glu Asp
            190                 195                 200

GAG AGG CTC ACC CCT CTC TGAAGGGCTG TTGTTCTGCT TCCTCAARAA              727
Glu Arg Leu Thr Pro Leu
        205

ATTAAACATT TGTTTCTGTG TGACTGCTGA GCATCCTGAA ATACCAAGAG CAGATCATAT     787

WTTTTGTTTC ACCATTCTTC TTTTGTAATA AATTTTGAAT GTGCTTGAAA AAAAAAAAAA     847

C                                                                     848
```

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 14 amino acids
- (B) TYPE: AMINO ACID
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
- (A) ORGANISM: Homo Sapiens

(ix) FEATURE:
- (A) NAME/KEY: sig_peptide
- (B) LOCATION: 1..14
- (C) IDENTIFICATION METHOD: Von Heijne matrix
- (D) OTHER INFORMATION: score 10.7
  seq LWLLFFLVTAIHA/EL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

```
Met Leu Trp Leu Leu Phe Phe Leu Val Thr Ala Ile His Ala
```

1 5 10

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

GGGAAGATGG AGATAGTATT GCCTG 25

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

CTGCCATGTA CATGATAGAG AGATTC 26

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 546 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:
(A) NAME/KEY: promoter
(B) LOCATION: 1..517

(ix) FEATURE:
(A) NAME/KEY: transcription start site
(B) LOCATION: 518

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: 17..25
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name CMYB_01

score 0.983
sequence TGTCAGTTG

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: complement(18..27)
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name MYOD_Q6
score 0.961
sequence CCCAACTGAC

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: complement(75..85)
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name S8_01
score 0.960
sequence AATAGAATTAG

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: 94..104
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name S8_01
score 0.966
sequence AACTAAATTAG

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: complement(129..139)
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name DELTAEF1_01
score 0.960
sequence GCACACCTCAG

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: complement(155..165)
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name GATA_C
score 0.964
sequence AGATAAATCCA

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: 170..178
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name CMYB_01
score 0.958
sequence CTTCAGTTG

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: 176..189
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name GATA1_02
score 0.959
sequence TTGTAGATAGGACA

(ix) FEATURE:
(A) NAME/KEY: TF binding-site

(B) LOCATION: 180..190
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name GATA_C
score 0.953
sequence AGATAGGACAT

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: 284..299
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name TAL1ALPHAE47_01
score 0.973
sequence CATAACAGATGGTAAG

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: 284..299
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name TAL1BETAE47_01
score 0.983
sequence CATAACAGATGGTAAG

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: 284..299
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name TAL1BETAITF2_01
score 0.978
sequence CATAACAGATGGTAAG

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: complement(287..296)
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name MYOD_Q6
score 0.954
sequence ACCATCTGTT

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: complement(302..314)
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name GATA1_04
score 0.953
sequence TCAAGATAAAGTA

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: 393..405
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name IK1_01
score 0.963
sequence AGTTGGGAATTCC

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: 393..404
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name IK2_01
score 0.985
sequence AGTTGGGAATTC

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: 396..405
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name CREL_01
score 0.962
sequence TGGGAATTCC

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: 423..436
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name GATA1_02
score 0.950
sequence TCAGTGATATGGCA

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: complement(478..489)
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name SRY_02
score 0.951
sequence TAAAACAAAACA

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: 486..493
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name E2F_02
score 0.957
sequence TTTAGCGC

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: complement(514..521)
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name MZF1_01
score 0.975
sequence TGAGGGGA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

TGAGTGCAGT GTTACATGTC AGTTGGGTTA AGTTTGTTAA TGTCATTCAA ATCTTCTATG 60

TCTTGATTTG CCTGCTAATT CTATTATTTC TGGAACTAAA TTAGTTTGAT GGTTCTATTA 120

GTTATTGACT GAGGTGTGCT AATCTCCCAT TATGTGGATT TATCTATTTC TTCAGTTGTA 180

GATAGGACAT TGATAGATAC ATAAGTACCA GGACAAAAGC AGGGAGATCT TTTTTCCAAA 240

ATCAGGAGAA AAAAATGACA TCTGGAAAAC CTATAGGGAA AGGCATAACA GATGGTAAGG 300

ATACTTTATC TTGAGTAGGA GAGCCTTCCT GTGGCAACGT GGAGAAGGGA AGAGGTCGTA 360

GAATTGAGGA GTCAGCTCAG TTAGAAGCAG GGAGTTGGGA ATTCCGTTCA TGTGATTTAG 420

CATCAGTGAT ATGGCAAATG TGGGACTAAG GGTAGTGATC AGAGGGTTAA AATTGTGTGT 480

TTTGTTTTAG CGCTGCTGGG GCATCGCCTT GGGTCCCCTC AAACAGATTC CCATGAATCT 540

CTTCAT 546

(2) INFORMATION FOR SEQ ID NO: 32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

GTACCAGGGA CTGTGACCAT TGC 23

(2) INFORMATION FOR SEQ ID NO: 33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

CTGTGACCAT TGCTCCCAAG AGAG 24

(2) INFORMATION FOR SEQ ID NO: 34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 861 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:
(A) NAME/KEY: promoter
(B) LOCATION: 1..806

(ix) FEATURE:
(A) NAME/KEY: transcription start site
(B) LOCATION: 807

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: complement(60..70)
(C) IDENTIFICATION METHOD: matinspector prediction

```
        (D) OTHER INFORMATION:  name NFY_Q6
                                score 0.956
                                sequence GGACCAATCAT

    (ix) FEATURE:
        (A) NAME/KEY: TF binding-site
        (B) LOCATION: 70..77
        (C) IDENTIFICATION METHOD: matinspector prediction
        (D) OTHER INFORMATION:  name MZF1_01
                                score 0.962
                                sequence CCTGGGGA

    (ix) FEATURE:
        (A) NAME/KEY: TF binding-site
        (B) LOCATION: 124..132
        (C) IDENTIFICATION METHOD: matinspector prediction
        (D) OTHER INFORMATION:  name CMYB_01
                                score 0.994
                                sequence TGACCGTTG

    (ix) FEATURE:
        (A) NAME/KEY: TF binding-site
        (B) LOCATION: complement(126..134)
        (C) IDENTIFICATION METHOD: matinspector prediction
        (D) OTHER INFORMATION:  name VMYB_02
                                score 0.985
                                sequence TCCAACGGT

    (ix) FEATURE:
        (A) NAME/KEY: TF binding-site
        (B) LOCATION: 135..143
        (C) IDENTIFICATION METHOD: matinspector prediction
        (D) OTHER INFORMATION:  name STAT_01
                                score 0.968
                                sequence TTCCTGGAA

    (ix) FEATURE:
        (A) NAME/KEY: TF binding-site
        (B) LOCATION: complement(135..143)
        (C) IDENTIFICATION METHOD: matinspector prediction
        (D) OTHER INFORMATION:  name STAT_01
                                score 0.951
                                sequence TTCCAGGAA

    (ix) FEATURE:
        (A) NAME/KEY: TF binding-site
        (B) LOCATION: complement(252..259)
        (C) IDENTIFICATION METHOD: matinspector prediction
        (D) OTHER INFORMATION:  name MZF1_01
                                score 0.956
                                sequence TTGGGGGA

    (ix) FEATURE:
        (A) NAME/KEY: TF binding-site
        (B) LOCATION: 357..368
        (C) IDENTIFICATION METHOD: matinspector prediction
        (D) OTHER INFORMATION:  name IK2_01
                                score 0.965
                                sequence GAATGGGATTTC

    (ix) FEATURE:
```

(A) NAME/KEY: TF binding-site
(B) LOCATION: 384..391
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name MZF1_01
score 0.986
sequence AGAGGGGA

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: complement(410..421)
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name SRY_02
score 0.955
sequence GAAAACAAAACA

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: 592..599
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name MZF1_01
score 0.960
sequence GAAGGGGA

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: 618..627
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name MYOD_Q6
score 0.981
sequence AGCATCTGCC

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: 632..642
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name DELTAEF1_01
score 0.958
sequence TCCCACCTTCC

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: complement(813..823)
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name S8_01
score 0.992
sequence GAGGCAATTAT

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: complement(824..831)
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name MZF1_01
score 0.986
sequence AGAGGGGA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

TACTATAGGG CACGCGTGGT CGACGGCCGG GCTGTTCTGG AGCAGAGGGC ATGTCAGTAA 60

TGATTGGTCC CTGGGGAAGG TCTGGCTGGC TCCAGCACAG TGAGGCATTT AGGTATCTCT 120

CGGTGACCGT TGGATTCCTG GAAGCAGTAG CTGTTCTGTT TGGATCTGGT AGGGACAGGG 180

CTCAGAGGGC TAGGCACGAG GGAAGGTCAG AGGAGAAGGS AGGSARGGCC CAGTGAGARG 240

GGAGCATGCC TTCCCCCAAC CCTGGCTTSC YCTTGGYMAM AGGGCGKTTY TGGGMACTTR 300

AAYTCAGGGC CCAASCAGAA SCACAGGCCC AKTCNTGGCT SMAAGCACAA TAGCCTGAAT 360

GGGATTTCAG GTTAGNCAGG GTGAGAGGGG AGGCTCTCTG GCTTAGTTTT GTTTTGTTTT 420

CCAAATCAAG GTAACTTGCT CCCTTCTGCT ACGGGCCTTG GTCTTGGCTT GTCCTCACCC 480

AGTCGGAACT CCCTACCACT TTCAGGAGAG TGGTTTTAGG CCCGTGGGGC TGTTCTGTTC 540

CAAGCAGTGT GAGAACATGG CTGGTAGAGG CTCTAGCTGT GTGCGGGGCC TGAAGGGGAG 600

TGGGTTCTCG CCCAAAGAGC ATCTGCCCAT TTCCCACCTT CCCTTCTCCC ACCAGAAGCT 660

TGCCTGAGCT GTTTGGACAA AAATCCAAAC CCCACTTGGC TACTCTGGCC TGGCTTCAGC 720

TTGGAACCCA ATACCTAGGC TTACAGGCCA TCCTGAGCCA GGGGCCTCTG GAAATTCTCT 780

TCCTGATGGT CCTTTAGGTT TGGGCACAAA ATATAATTGC CTCTCCCCTC TCCCATTTTC 840

TCTCTTGGGA GCAATGGTCA C 861

(2) INFORMATION FOR SEQ ID NO: 35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

CTGGGATGGA AGGCACGGTA 20

(2) INFORMATION FOR SEQ ID NO: 36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

GAGACCACAC AGCTAGACAA 20

(2) INFORMATION FOR SEQ ID NO: 37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 555 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:
(A) NAME/KEY: promoter
(B) LOCATION: 1..500

(ix) FEATURE:
(A) NAME/KEY: transcription start site
(B) LOCATION: 501

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: 191..206
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name ARNT_01
score 0.964
sequence GGACTCACGTGCTGCT

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: 193..204
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name NMYC_01
score 0.965
sequence ACTCACGTGCTG

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: 193..204
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name USF_01
score 0.985
sequence ACTCACGTGCTG

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: complement(193..204)
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name USF_01
score 0.985
sequence CAGCACGTGAGT

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: complement(193..204)
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name NMYC_01
score 0.956
sequence CAGCACGTGAGT

(ix) FEATURE:

(A) NAME/KEY: TF binding-site
(B) LOCATION: complement(193..204)
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name MYCMAX_02
score 0.972
sequence CAGCACGTGAGT

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: 195..202
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name USF_C
score 0.997
sequence TCACGTGC

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: complement(195..202)
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name USF_C
score 0.991
sequence GCACGTGA

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: complement(210..217)
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name MZF1_01
score 0.968
sequence CATGGGGA

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: 397..410
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name ELK1_02
score 0.963
sequence CTCTCCGGAAGCCT

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: 400..409
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name CETS1P54_01
score 0.974
sequence TCCGGAAGCC

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: complement(460..470)
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name AP1_Q4
score 0.963
sequence AGTGACTGAAC

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: complement(460..470)
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name AP1FJ_Q2
score 0.961

sequence AGTGACTGAAC

(ix) FEATURE:
(A) NAME/KEY: TF binding-site
(B) LOCATION: 547..555
(C) IDENTIFICATION METHOD: matinspector prediction
(D) OTHER INFORMATION: name PADS_C
score 1.000
sequence TGTGGTCTC

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

```
CTATAGGGCA CGCKTGGTCG ACGGCCCGGG CTGGTCTGGT CTGTKGTGGA GTCGGGTTGA   60
AGGACAGCAT TTGTKACATC TGGTCTACTG CACCTTCCCT CTGCCGTGCA CTTGGCCTTT  120
KAWAAGCTCA GCACCGGTGC CCATCACAGG GCCGGCAGCA CACACATCCC ATTACTCAGA  180
AGGAACTGAC GGACTCACGT GCTGCTCCGT CCCCATGAGC TCAGTGGACC TGTCTATGTA  240
GAGCAGTCAG ACAGTGCCTG GGATAGAGTG AGAGTTCAGC CAGTAAATCC AAGTGATTGT  300
CATTCCTGTC TGCATTAGTA ACTCCCAACC TAGATGTGAA AACTTAGTTC TTTCTCATAG  360
GTTGCTCTGC CCATGGTCCC ACTGCAGACC CAGGCACTCT CCGGAAGCCT GGAAATCACC  420
CGTGTCTTCT GCCTGCTCCC GCTCACATCC CACACTTGTG TTCAGTCACT GAGTTACAGA  480
TTTTGCCTCC TCAATTTCTC TTGTCTTAGT CCCATCCTCT GTTCCCCTGG CCAGTTTGTC  540
TAGCTGTGTG GTCTC                                                   555
```

(2) INFORMATION FOR SEQ ID NO: 38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 195 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 109..171
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 14.3
seq LLLCAVLLSLASA/SS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

```
AGAGACTAGA GCAGGAAGAG CAGCGGCGAG GCGGCGGTGG TGGCTGAGTC CGTGGTGGCA   60
```

```
GAGGCGAAGG CGACAGCTCT AGGGGTTGGC ACCGGCCCCG AGAGGAGG ATG CGG GTC    117
                                                     Met Arg Val
                                                         -20

CGG ATA GGG CTG ACG CTG CTG CTG TGT GCG GTG CTG CTG AGC TTG GCC    165
Arg Ile Gly Leu Thr Leu Leu Leu Cys Ala Val Leu Leu Ser Leu Ala
            -15                 -10                 -5

TCG GCG TCC TCG GAT GAA GAA GGC AAT GGG                            195
Ser Ala Ser Ser Asp Glu Glu Gly Asn Gly
         1               5
```

(2) INFORMATION FOR SEQ ID NO: 39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 162 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 55..138
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 11.1
seq GLLFLLLLLMLLA/DP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

```
AAGAGCCCAG AGAGCTGAAC CTGCATCCCG GACCTGCGGC GACCGTCGTA CACC ATG     57
                                                            Met

GGC CTC CAC CTC CGC CCC TAC CGT GTG GGG CTG CTC CCG GAT GGC CTC    105
Gly Leu His Leu Arg Pro Tyr Arg Val Gly Leu Leu Pro Asp Gly Leu
        -25                 -20                 -15

CTG TTC CTC TTG CTG CTG CTA ATG CTG CTC GCG GAC CCA GCG CTC CCG    153
Leu Phe Leu Leu Leu Leu Leu Met Leu Leu Ala Asp Pro Ala Leu Pro
    -10                 -5                   1               5

GCC GCT AGG                                                        162
Ala Ala Arg
```

(2) INFORMATION FOR SEQ ID NO: 40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 421 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 20..82
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 10.6
seq LLLGAVSWPPASA/SG

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

```
AGCGTGGCCG CCCGGGGCC ATG GCG ACA CTC AGC TTC GTC TTC CTG CTG CTG     52
                     Met Ala Thr Leu Ser Phe Val Phe Leu Leu Leu
                         -20                 -15

GGG GCA GTG TCC TGG CCT CCG GCT TCT GCC TCC GGC CAG GAG TTC TGG     100
Gly Ala Val Ser Trp Pro Pro Ala Ser Ala Ser Gly Gln Glu Phe Trp
-10                 -5                   1               5

CCC GGA CAA TCG GCG GCC GAT ATT CTG TCG GGG GCG GCT TCC CGC AGA     148
Pro Gly Gln Ser Ala Ala Asp Ile Leu Ser Gly Ala Ala Ser Arg Arg
            10                  15                  20

CGG TAT CTT CTG TAT GAC GTC AAC CCC CCG GAA GGC TTC AAC CTG CGC     196
Arg Tyr Leu Leu Tyr Asp Val Asn Pro Pro Glu Gly Phe Asn Leu Arg
        25                  30                  35

AGG GAT GTC TAT ATC CGA ATC GCC TCT CTC CTG AAG ACT CTG CTG AAG     244
Arg Asp Val Tyr Ile Arg Ile Ala Ser Leu Leu Lys Thr Leu Leu Lys
    40                  45                  50

ACG GAG GAG TGG GTG CTT RTC CTG CCT CCA TGG GGC CGC CTC TNN RAC     292
Thr Glu Glu Trp Val Leu Xaa Leu Pro Pro Trp Gly Arg Leu Xaa Xaa
 55                  60                  65                  70

TGG CAG AGT SST GAC ATC CAC CAG GTC CGG ATT CCC TGG TCT GAG TTT     340
Trp Gln Ser Xaa Asp Ile His Gln Val Arg Ile Pro Trp Ser Glu Phe
                75                  80                  85

TTT GAT CTT CCA AGT CTC AAT AAA AAC ATC CCC GTC ATC GAG TAT GAG     388
Phe Asp Leu Pro Ser Leu Asn Lys Asn Ile Pro Val Ile Glu Tyr Glu
            90                  95                 100

CAG TTC ATC GCA GAA TCT GGT GGG CCC TTT ATT                         421
Gln Phe Ile Ala Glu Ser Gly Gly Pro Phe Ile
        105                 110
```

(2) INFORMATION FOR SEQ ID NO: 41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 182 base pairs

```
        (B) TYPE: NUCLEIC ACID
        (C) STRANDEDNESS: DOUBLE
        (D) TOPOLOGY: LINEAR

    (ii) MOLECULE TYPE: CDNA

    (vi) ORIGINAL SOURCE:
        (A) ORGANISM: Homo Sapiens
        (D) DEVELOPMENTAL STAGE: Fetal
        (F) TISSUE TYPE: brain

    (ix) FEATURE:
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: 6..167
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 9.1
                                seq VLCLRGLVSLAFQ/GP

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

AACAA ATG TTT TTA TTT TTG TCA CCA GCC ACC CCT GTC CTG CCG CCT TCT    50
      Met Phe Leu Phe Leu Ser Pro Ala Thr Pro Val Leu Pro Pro Ser
                      -50                 -45                 -40

CTC GAC TCC AGA GAC CTG TTG CCT CAT CTC TTT TGG GGA AGA GCC GGC      98
Leu Asp Ser Arg Asp Leu Leu Pro His Leu Phe Trp Gly Arg Ala Gly
                -35                 -30                 -25

AGC TCC TCC TCA TCC CCT GCC TTA AGT CCA GTT CTT TGC CTC AGG GGT     146
Ser Ser Ser Ser Ser Pro Ala Leu Ser Pro Val Leu Cys Leu Arg Gly
            -20                 -15                 -10

CTC GTT TCC TTG GCC TTC CAG GGT CCC CAC CCC GAG                     182
Leu Val Ser Leu Ala Phe Gln Gly Pro His Pro Glu
        -5                  1               5


(2) INFORMATION FOR SEQ ID NO: 42:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 272 base pairs
        (B) TYPE: NUCLEIC ACID
        (C) STRANDEDNESS: DOUBLE
        (D) TOPOLOGY: LINEAR

    (ii) MOLECULE TYPE: CDNA

    (vi) ORIGINAL SOURCE:
        (A) ORGANISM: Homo Sapiens
        (F) TISSUE TYPE: Substantia nigra

    (ix) FEATURE:
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: 63..125
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 8.8
                                seq LLWALLFMQSLWP/QL

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:
```

```
AGCAGTGCAG CATTAATGGG CCGCTGACAT GAATATGGAG TAGTTTTCTC TAGCAAAGAG   60

TA ATG TGG GCC ATG GAG TCA GGC CAC CTC CTC TGG GCT CTG CTG TTC     107
   Met Trp Ala Met Glu Ser Gly His Leu Leu Trp Ala Leu Leu Phe
       -20                 -15                 -10

ATG CAG TCC TTG TGG CCT CAA CTG ACT GAT GGA GCC ACT CGA GTC TAC    155
Met Gln Ser Leu Trp Pro Gln Leu Thr Asp Gly Ala Thr Arg Val Tyr
     -5                  1               5                  10

TAC CTG GGC ATC CGG GAT GTG CAG TGG AAC TAT GCT CCC AAG GGA AGA    203
Tyr Leu Gly Ile Arg Asp Val Gln Trp Asn Tyr Ala Pro Lys Gly Arg
                 15                  20                  25

AAT GTC ATC ACG AAC CAG CCT CTG GAC AGT GAC ATA GTG GCT TCC AGC    251
Asn Val Ile Thr Asn Gln Pro Leu Asp Ser Asp Ile Val Ala Ser Ser
             30                  35                  40

TTC TTA AAG TCT GAC GAT GGG                                        272
Phe Leu Lys Ser Asp Asp Gly
         45
```

(2) INFORMATION FOR SEQ ID NO: 43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 195 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 130..186
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.7
seq ILGLLCCVLATMA/NP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

```
AAACTAAAAT TTCATGCAAA GTTGATGCCA TTTCAACATG AATATGATTC TATATAGTAA   60

GTTGAATGTG TAGAAATTCA GATATCTGGA AGTTATGAAA CAAGGAATCA ATTAGGTATA  120

CAAAAGCCG ATG ACT CAC TAT AGA AAT ATT CTT GGT CTC CTG TGC TGT GTA  171
          Met Thr His Tyr Arg Asn Ile Leu Gly Leu Leu Cys Cys Val
                          -15                 -10

TTA GCA ACC ATG GCC AAC CCC GGG                                    195
Leu Ala Thr Met Ala Asn Pro Gly
```

-5 1

(2) INFORMATION FOR SEQ ID NO: 44:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 333 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 253..297
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.4
seq LLLLLASLIERSS/KT

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

```
AANTCATTCT TTTGCCTGGA GTTTTGTGAG GTACCCGCTT GCTTTATGGG AAAAGGCTGC   60

TCCGGAACTG CCCTACTTTA GACTTTTTCA TGGTTATCAA TCKGKACAMA GAATCACCAA  120

ACTGATAAAG CAGGAACNAG AGGGCAAATC ACGCTGCCAA GACAACTGTG TAATTCGCTC  180

GAAAAAGAAA CGAAGACAAT GTATATAAAA ATATGCAAGA ATCACAGGAA ACCCACATAT  240

CCAACCACCT AG ATG AAG TTG TTG CTG CTG TTA GCA TCA CTC ATA GAA AGA  291
              Met Lys Leu Leu Leu Leu Leu Ala Ser Leu Ile Glu Arg
              -15                 -10                     -5

AGT TCC AAA ACA AGC TGC TWN NGA CAG CAC TAT TCC AGC CAG           333
Ser Ser Lys Thr Ser Cys Xaa Xaa Gln His Tyr Ser Ser Gln
         1               5                   10
```

(2) INFORMATION FOR SEQ ID NO: 45:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 219 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:

(A) NAME/KEY: sig_peptide
(B) LOCATION: 64..177
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.2
seq SFXLFLALCASFS/FF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

```
ATATTTAATT TATTATCTTC CTTAATTTGT AGATGAGAAA CTGAGGCATA GTAATGGACA    60

GGA ATG GCT AGG AAT CAA GCC TTG GTC TGC CTT CCC TCT TTT CAG AAT     108
    Met Ala Arg Asn Gln Ala Leu Val Cys Leu Pro Ser Phe Gln Asn
                -35                 -30                 -25

GCC TTT ATT CCT GTT GAA GAT CTT CCC ACA TCA TTT TKG CTC TTT CTC     156
Ala Phe Ile Pro Val Glu Asp Leu Pro Thr Ser Phe Xaa Leu Phe Leu
            -20                 -15                 -10

GCC CTC TGT GCT TCC TTC TCC TTT TTT TTA TKT CTT TCT CTT TCC CTC     204
Ala Leu Cys Ala Ser Phe Ser Phe Phe Leu Xaa Leu Ser Leu Ser Leu
        -5                   1                   5

CCT TCC TTC TTT TTT                                                 219
Pro Ser Phe Phe Phe
 10
```

(2) INFORMATION FOR SEQ ID NO: 46:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 286 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 164..268
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.2
seq SLLLLFYSFYVLA/VK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

```
ACATTATTTG TCAAAACCTC ATTAAGGTTT TGACAACAAA TGATAATTGG AAACAGTTTT    60

CTGGCTCCTT TAGTAGTAGC ATGTCATTTT ATCTGCTTCT TGGCCTTAGT ATGTTTGAAG   120

TGAAACTGTT WRABVWAGAT GTTTGTATTT GGTTCAAGCT TAA ATG CCA AAT GAA     175
                                                Met Pro Asn Glu
                                                -35
```

```
TCT TGG CAA ATT CCA TGT GGC AAG CAA GAA GCT GAA ACT CTA TTC AAC      223
Ser Trp Gln Ile Pro Cys Gly Lys Gln Glu Ala Glu Thr Leu Phe Asn
    -30                 -25                 -20

TTC CAA AGC TTG TTG TTG TTA TTT TAT TCC TTC TAT GTT CTA GCT GTT      271
Phe Gln Ser Leu Leu Leu Leu Phe Tyr Ser Phe Tyr Val Leu Ala Val
-15                 -10                 -5                    1

AAA AGA GGG CCA GGG                                                  286
Lys Arg Gly Pro Gly
            5
```

(2) INFORMATION FOR SEQ ID NO: 47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 209 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 93..200
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8
seq LVLLICLVSSYLP/QL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

```
AATTACTTCA ACTGTTGACT ATATTTTGTT TATRTTTCTC ATACTGTTGT TTCTTGAATC     60

CATTCCAGTC AGGTGTGTGC CCCCATCACT TC ATG CAA ACT ACT TTT ATC GAT      113
                                    Met Gln Thr Thr Phe Ile Asp
                                        -35                 -30

GTC ACT GTT GAC CAA CAT GTT GCT AAA TCT AAT GAT CAC TTG TCA GTC      161
Val Thr Val Asp Gln His Val Ala Lys Ser Asn Asp His Leu Ser Val
                -25                 -20                 -15

CTA GTC TTA CTT ATT TGC TTA GTC TCA TCT TAT TTG CCA CAG CTC CCG      209
Leu Val Leu Leu Ile Cys Leu Val Ser Ser Tyr Leu Pro Gln Leu Pro
            -10                 -5                   1
```

(2) INFORMATION FOR SEQ ID NO: 48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 363 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 124..342
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.8
seq YLPLLAGLGLTLA/AP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

```
AAGGGTCCGC CAGCTCGGGG CCAGCGCATG GGGCTGCTGA GACCGCTCCG GACGTGCGAA   60

SGTTCGCGGT GCGGTAGGAA GCTGAAGTCT TCAGCAGAAC CCTGACCTAG GTAGGGAGTC  120

TCA ATG GCC CTG GGT GAA GAA AAG GCA GAA GCG GAA GCA TCT GRR GAC     168
    Met Ala Leu Gly Glu Glu Lys Ala Glu Ala Glu Ala Ser Xaa Asp
                -70                 -65                 -60

ACA AAG GCC CAG TCC TAT GGG AGA GGG AGC TGC AGG GAG CGG GAG CTG     216
Thr Lys Ala Gln Ser Tyr Gly Arg Gly Ser Cys Arg Glu Arg Glu Leu
            -55                 -50                 -45

GAC ATC CCA GGG CCC ATG AGT GGG GAG CAG CCC CCA CGC CTG GAA GCT     264
Asp Ile Pro Gly Pro Met Ser Gly Glu Gln Pro Pro Arg Leu Glu Ala
        -40                 -35                 -30

GAG GGA GGG CTC ATC TCC CCT GTA TGG GGG GCA GAA NGA TAC CTG CCC     312
Glu Gly Gly Leu Ile Ser Pro Val Trp Gly Ala Glu Xaa Tyr Leu Pro
    -25                 -20                 -15

CTA CTT GCT GGA TTG GGA CTG ACC CTG GCG GCC CCT CTA GAG CCC ACA     360
Leu Leu Ala Gly Leu Gly Leu Thr Leu Ala Ala Pro Leu Glu Pro Thr
-10                  -5                   1               5

ACG                                                                 363
Thr
```

(2) INFORMATION FOR SEQ ID NO: 49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 271 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 149..211
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.7
seq LLCISPFVPFTSG/NK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

```
AAAACTCTAT GAGTGTCTTT TTGAGACCAT AAAGCAGACT TTAGTAACTT TCTATTTCTG    60

TAAGTACTAA ATGTCTGGCA TTTTAAACTT TTGTAGAATA CATAATGTTG RACACTGGAA   120

TAATACTATT KATTTTCACC TGTGAAAA ATG ACT TCA TTG TAC TTG AAA CAC      172
                               Met Thr Ser Leu Tyr Leu Lys His
                                   -20                 -15

CTC CTT TGC ATT TCT CCA TTT GTG CCA TTC ACT AGT GGA AAT AAA TTG     220
Leu Leu Cys Ile Ser Pro Phe Val Pro Phe Thr Ser Gly Asn Lys Leu
            -10                 -5                   1

TAT TAT ACC ATG ATC TAC TGG CTT TTT AAA ACT GTA TTA AAT ATG CAC     268
Tyr Tyr Thr Met Ile Tyr Trp Leu Phe Lys Thr Val Leu Asn Met His
     5                   10                  15

GGG                                                                 271
Gly
 20
```

(2) INFORMATION FOR SEQ ID NO: 50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 378 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 205..345
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.6
seq CLATLTLFHTSFS/FQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:

```
AGGGCTCTCC ATGAAGCGGG TGTTGTCAAC TTACTTGGAA CCAGATATTG AGCACTTATT    60

TTATATCAGG CAGAATACAG AGCCACTTAY ATGRKWTCAA TKAATTCTTA CAACTCCCTT   120

GGGAARGAGG TGTTTGTCTT CTTTCCATTT AACAGATGAG AAGACTGAGA CTTGATGAGT   180
```

```
TTAGTCAATT GTTCTGAGTA GTAA ATG ACA GAC TCT CCC AAT GCT CAT GGC      231
                           Met Thr Asp Ser Pro Asn Ala His Gly
                                   -45                 -40

TTA GCT CTC ACC ACC AAG TGG ATG ATG CCT GCT GTC TCT TTG AAC TTG     279
Leu Ala Leu Thr Thr Lys Trp Met Met Pro Ala Val Ser Leu Asn Leu
            -35                 -30                 -25

ACC TAT TAC TTG CCA TCT TGG TAC CTT TGT TTG GCC ACT CTT ACT TTA     327
Thr Tyr Tyr Leu Pro Ser Trp Tyr Leu Cys Leu Ala Thr Leu Thr Leu
        -20                 -15                 -10

TTC CAC ACC TCT TTC TCC TTC CAA GCT TCT GAG TCT GTC AAA GCC ATC     375
Phe His Thr Ser Phe Ser Phe Gln Ala Ser Glu Ser Val Lys Ala Ile
    -5                  1               5                   10

ACG                                                                 378
Thr
```

(2) INFORMATION FOR SEQ ID NO: 51:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 376 base pairs
 (B) TYPE: NUCLEIC ACID
 (C) STRANDEDNESS: DOUBLE
 (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: Homo Sapiens
 (F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
 (A) NAME/KEY: sig_peptide
 (B) LOCATION: 143..286
 (C) IDENTIFICATION METHOD: Von Heijne matrix
 (D) OTHER INFORMATION: score 7.5
 seq FVILLLFIFTVVS/LV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:

```
AATCGCTTCA GCAGCATCCT CTCAGACAAG AGCCACTATT TCTGATTCAG ATCACCTGTC    60

ATCGAAGTTT AAAGAAGGGG AAACAGGAGA CAGAAATACA CTGAACCAAA AAGATTCAAA   120

AGAGCAAGTG GAATCTCTAA GA ATG GCT TCC AGC CAC TGG AAT GAA ACC ACT    172
                         Met Ala Ser Ser His Trp Asn Glu Thr Thr
                                     -45                 -40

ACC TCT GTT TAT CAG TAC CTT GGT TTT CAA GTT CAA AAA ATT TAC CCT     220
Thr Ser Val Tyr Gln Tyr Leu Gly Phe Gln Val Gln Lys Ile Tyr Pro
            -35                 -30                 -25

TTC CAT GAC AAC TGG AAC ACT GCC TGC TTT GTC ATC CTG CTT TTA TTT     268
Phe His Asp Asn Trp Asn Thr Ala Cys Phe Val Ile Leu Leu Leu Phe
        -20                 -15                 -10
```

```
ATA TTT ACA GTG GTA TCT TTA GTG GTG CTG GCT TTC CTT TAT GAA GTG      316
Ile Phe Thr Val Val Ser Leu Val Val Leu Ala Phe Leu Tyr Glu Val
     -5                   1               5                   10

CTT GAM WGC TGC TGC TGT GTA AAA AAC AAA ACC GTG AAA GAC TTG AAA      364
Leu Xaa Xaa Cys Cys Cys Val Lys Asn Lys Thr Val Lys Asp Leu Lys
                 15                  20                  25

AGT GAA CCC AAG                                                      376
Ser Glu Pro Lys
            30
```

(2) INFORMATION FOR SEQ ID NO: 52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 429 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 298..372
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.4
seq IFLLNMWVACLLS/GE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:

```
ACTGAGATCC ACCTGGGCTC CCAAGACAGA TAGTGGGGTC TTGTCCAGTT CATATATTAT     60

CTACAGTTGC CAATGTGATA TAAAATATGA TTTAGGTCAT ATCACCTACT TGAAAGTCTG    120

CACTGGTGCT GTCTTGGTAA CAGAAAGAGA AAGACTGAAT TCCCTTATTT TGCCCATCAG    180

GCTCTCCTTG ATATAGTCTC TGCCCACCCA CCTCTCCAAC CTCACATCAT CTCTCTTCTG    240

CCTCATACGC TATGCTCCGG CACGTATAGG TTCCTATACA ATTTTGTTTC ATACTTG       297

ATG TCT TTA CTT TTT GTC TTT TGC CTG GAA TGC AGT ATT TTT CTA TTG      345
Met Ser Leu Leu Phe Val Phe Cys Leu Glu Cys Ser Ile Phe Leu Leu
-25                 -20                 -15                 -10

AAT ATG TGG GTT GCT TGC CTT CTG AGT GGT GAG ATT CCC CAT TCC TCA      393
Asn Met Trp Val Ala Cys Leu Leu Ser Gly Glu Ile Pro His Ser Ser
                 -5                   1                   5

TGG AAN MTA AAG TTA ATT GGC ACC TTG CCC ACT TCT                      429
Trp Xaa Xaa Lys Leu Ile Gly Thr Leu Pro Thr Ser
         10                  15
```

(2) INFORMATION FOR SEQ ID NO: 53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 244 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 149..202
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.4
seq VLLLLPLVAFITL/KF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

```
AAAATTTCAA AACGTAGGAT AAAAGGAATC AAATGTATTA ATGGAWAAKC AACTGAACTT    60

AATTTCGATT CTTTTCTATC ATTTTTTCCT AGGCTAKAGA TAGACTAAAT TCATATCTGA   120

AAATTCTCAA TTTTTGAGAA AAGACAAA ATG TTT GTC GTT ACA GTT TTG TTG     172
                               Met Phe Val Val Thr Val Leu Leu
                                                -15

TTG CTG CCC TTA GTT GCT TTC ATT ACC CTC AAA TTC TGT AAC TTG ATT    220
Leu Leu Pro Leu Val Ala Phe Ile Thr Leu Lys Phe Cys Asn Leu Ile
-10                 -5                      1               5

AAT TTT CCA ACT CWK AGA CAC GGG                                    244
Asn Phe Pro Thr Xaa Arg His Gly
            10
```

(2) INFORMATION FOR SEQ ID NO: 54:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 212 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide

(B) LOCATION: 12..77
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.2
seq IIYALQFLFLVFA/PS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

```
AATTCAATCA A ATG AAC CGA TCT TGT AGA AAC ACT GGA ATC ATT TAT GCG      50
             Met Asn Arg Ser Cys Arg Asn Thr Gly Ile Ile Tyr Ala
                     -20                 -15                 -10

TTG CAG TTT CTC TTT CTT GTT TTC GCT CCT TCT TCC CTG GGA TAT TTT       98
Leu Gln Phe Leu Phe Leu Val Phe Ala Pro Ser Ser Leu Gly Tyr Phe
                 -5                   1               5

GAG TGG ATT GTG GCT ATT AAT CAA GAT CTC GTG CTA TTC GTG TTT TGC      146
Glu Trp Ile Val Ala Ile Asn Gln Asp Leu Val Leu Phe Val Phe Cys
         10                  15                  20

TTG TCG TTT TCG CTC AGG ATT AGC ATC ATT CAA GGC AAA CGC AAA GCT      194
Leu Ser Phe Ser Leu Arg Ile Ser Ile Ile Gln Gly Lys Arg Lys Ala
     25                  30                  35

GCT TTT CCC ACC CCC CCC                                              212
Ala Phe Pro Thr Pro Pro
 40                  45
```

(2) INFORMATION FOR SEQ ID NO: 55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 221 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 87..191
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.2
seq LSLLLAWVTLTHL/LS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:

```
ACCTTCGACC CCCACCCTGG GAATACCAAT CCAGTGATTC CACCATCTAC TCACTGTCCC      60

TCATCGTTGG TGTCTTCTCT AACTTC ATG ACC CAA ACC ACA TGG GGA GCC CCC     113
                             Met Thr Gln Thr Thr Trp Gly Ala Pro
                             -35                 -30

ACC AGG GCC AGC AAT CAC CCT CTC CCT GCA TGG CTC ACC CTC AGC CTC      161
```

```
Thr Arg Ala Ser Asn His Pro Leu Pro Ala Trp Leu Thr Leu Ser Leu
    -25                 -20                 -15

CTC CTG GCC TGG GTG ACC CTT ACA CAC CTT CTC TCT GTG CTC ACA CAT      209
Leu Leu Ala Trp Val Thr Leu Thr His Leu Leu Ser Val Leu Thr His
-10                 -5                   1                   5

CCA ACC CTC CTG                                                      221
Pro Thr Leu Leu
             10
```

(2) INFORMATION FOR SEQ ID NO: 56:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 328 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 74..121
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.1
seq XXAVLCVCAAAWC/SQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 56:

```
AGAAASNCCG AGTTGGCAGA GCAGGGCTGC ATTTCCAGCA GGAGCTGCGA GCACAGTGCT    60

GGCTCACAAC AAG ATG CTC AAG GDV NNA GCC GTA CTG TGT GTG TGT GCA      109
               Met Leu Lys Xaa Xaa Ala Val Leu Cys Val Cys Ala
                   -15                 -10                 -5

GCC GCT TGG TGC AGT CAG TCT CTC GCA GCT GCC GCG GCG GTG GCT GCA      157
Ala Ala Trp Cys Ser Gln Ser Leu Ala Ala Ala Ala Ala Val Ala Ala
                 1               5                  10

DCC GGG GGG CGG TCG GAC GGC GGT AAT TTT CTG GAT GAT AAA CAA TGG      205
Xaa Gly Gly Arg Ser Asp Gly Gly Asn Phe Leu Asp Asp Lys Gln Trp
        15                  20                  25

CTC ACC ACA ATC TCT CAG TAT GAC AAG GAA GTC GGA CAG TGG AAC AAA      253
Leu Thr Thr Ile Ser Gln Tyr Asp Lys Glu Val Gly Gln Trp Asn Lys
    30                  35                  40

TTC CGA GAC GAT GAT TAT TTC CGC ACT TGG AGT CCA GGA AAA CCC TTC      301
Phe Arg Asp Asp Asp Tyr Phe Arg Thr Trp Ser Pro Gly Lys Pro Phe
 45                  50                  55                  60

GAT CAG GCT TTA GAT CSA GCT AAC GGG                                  328
Asp Gln Ala Leu Asp Xaa Ala Asn Gly
```

65

(2) INFORMATION FOR SEQ ID NO: 57:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 233 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 168..227
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.9
seq LHLLGSSISPASA/SL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:

```
AATGAAAAGC AGTTTGTTAT TATGCAGGAA AATCAGTTTC ATCATTTTAG TTACACTAAA    60

CACTTTTGGC AGCTTAATAT GACCTTTTTA AATTTTTTKT TATTTTTTTT ATTTTTATTT   120

CTTTAAGATG GAGTCTTGCT CTGTTGCCCG GGCTGGAGTA CAATGGC ATG ATC TCA     176
                                                    Met Ile Ser
                                                    -20

GCT CAC TGC AAC CTC CAC CTC CTG GGT TCA AGC ATT TCT CCT GCC TCA     224
Ala His Cys Asn Leu His Leu Leu Gly Ser Ser Ile Ser Pro Ala Ser
        -15                 -10                 -5

GCC TCC CTG                                                         233
Ala Ser Leu
    1
```

(2) INFORMATION FOR SEQ ID NO: 58:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 170 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

```
(ix) FEATURE:
      (A) NAME/KEY: sig_peptide
      (B) LOCATION: 6..89
      (C) IDENTIFICATION METHOD: Von Heijne matrix
      (D) OTHER INFORMATION:  score 6.9
                              seq FLPFLLSLPLDQT/LP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

AAGAT ATG TKC SCB AAA GCC TGC AGA ACC CTC GCT TGG TTG CCT GAS CCG    50
      Met Xaa Xaa Lys Ala Cys Arg Thr Leu Ala Trp Leu Pro Xaa Pro
                  -25                     -20                 -15

TTC TTA CCC TTT CTC CTC AGT CTT CCC TTG GAC CAG ACG CTT CCT CGC      98
Phe Leu Pro Phe Leu Leu Ser Leu Pro Leu Asp Gln Thr Leu Pro Arg
            -10                 -5                  1

CAG GGT CCT GGC CAA TCC CTG TCC TTC CCA GAA AAC TAC CAG ACT CTT     146
Gln Gly Pro Gly Gln Ser Leu Ser Phe Pro Glu Asn Tyr Gln Thr Leu
    5                   10                  15

CCC AAG AGC ACC CGA CAC CCT GGG                                     170
Pro Lys Ser Thr Arg His Pro Gly
 20                 25


(2) INFORMATION FOR SEQ ID NO: 59:

  (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 192 base pairs
        (B) TYPE: NUCLEIC ACID
        (C) STRANDEDNESS: DOUBLE
        (D) TOPOLOGY: LINEAR

  (ii) MOLECULE TYPE: CDNA

  (vi) ORIGINAL SOURCE:
        (A) ORGANISM: Homo Sapiens
        (F) TISSUE TYPE: Substantia nigra

  (ix) FEATURE:
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: 19..75
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 6.8
                                seq GLLLVFLPHPQRG/GQ

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 59:

ACATCGCCCG AGCAGGGG ATG GCG GTG AAG CGG CTA GGG CTG CTG TTG GTG      51
                    Met Ala Val Lys Arg Leu Gly Leu Leu Leu Val
                                    -15                 -10

TTC CTG CCT CAT CCG CAG CGG GGA GGA CAG GAG AGG TCT GCC CAC ACC      99
Phe Leu Pro His Pro Gln Arg Gly Gly Gln Glu Arg Ser Ala His Thr
            -5                  1                   5

CCG AGG CAG CAC CCA GCT CGC CCC ACT TCC CTC TCG CAG GGG GAG AGA     147
```

```
Pro Arg Gln His Pro Ala Arg Pro Thr Ser Leu Ser Gln Gly Glu Arg
     10                  15                  20

CCA GGA CGC GGT GGG GGG TGG GGG AAT GGC CGT GAC GCC CAC CAG         192
Pro Gly Arg Gly Gly Gly Trp Gly Asn Gly Arg Asp Ala His Gln
 25                  30                  35
```

(2) INFORMATION FOR SEQ ID NO: 60:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 433 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 77..325
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.8
seq LLMILTFPFKILS/DA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:

```
AAGGCTGCAA TAACTACTAC TTACTGGATA CATTCAAACC CTCCAGAATC AACAGTTATC    60

AGGTAACCAA CAAGAA ATG CAA GCC GTC GAC AAC CTC ACC TCT GCG CCT GGG  112
                  Met Gln Ala Val Asp Asn Leu Thr Ser Ala Pro Gly
                              -80                 -75

AAC ACC AGT CTG TGC ACC AGA GAC TAC AAA ATC ACC CAG GTC CTC TTC     160
Asn Thr Ser Leu Cys Thr Arg Asp Tyr Lys Ile Thr Gln Val Leu Phe
    -70                 -65                 -60

CCA CTG CTC TAC ACT GTC CTG TTT TTT GTT GGA CTT ATC ACA AAT GGC     208
Pro Leu Leu Tyr Thr Val Leu Phe Phe Val Gly Leu Ile Thr Asn Gly
-55                 -50                 -45                 -40

CTG GCG ATG AGG ATT TTC TTT CAA ATC CGG AGT AAA TCA AAC TTT ATT     256
Leu Ala Met Arg Ile Phe Phe Gln Ile Arg Ser Lys Ser Asn Phe Ile
                -35                 -30                 -25

ATT TTT CTT AAG AAC ACA GTC ATT TCT GAT CTT CTC ATG ATT CTG ACT     304
Ile Phe Leu Lys Asn Thr Val Ile Ser Asp Leu Leu Met Ile Leu Thr
            -20                 -15                 -10

TTT CCA TTC AAA ATT CTT AGT GAT GCC AAA CTG GGA ACA GGA CCA CTG     352
Phe Pro Phe Lys Ile Leu Ser Asp Ala Lys Leu Gly Thr Gly Pro Leu
        -5                   1               5

AGA ACT TTT GTG TGT CAA GTT ACC TCC GTC ATA TTT TAT KKC RSA ATG     400
Arg Thr Phe Val Cys Gln Val Thr Ser Val Ile Phe Tyr Xaa Xaa Met
```

```
 10                  15                  20                  25

TAT ATC AGT ATT TCA TTC CTG GGA CTG ATA ACT                          433
Tyr Ile Ser Ile Ser Phe Leu Gly Leu Ile Thr
                 30                  35
```

(2) INFORMATION FOR SEQ ID NO: 61:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 176 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 12..119
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.8
seq LWFFLPSLXCPEC/CP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:

```
ATCTGGTCTG C ATG TGG ACC CTC CCC AGC CTC AGT GCA TCG TTT CAG CCT     50
             Met Trp Thr Leu Pro Ser Leu Ser Ala Ser Phe Gln Pro
                 -35                 -30                 -25

TTT CTG GGC AGC CTT CGC CCT TCT CAC ATC CTG TGG TTC TTC CTG CCC      98
Phe Leu Gly Ser Leu Arg Pro Ser His Ile Leu Trp Phe Phe Leu Pro
            -20                 -15                 -10

TCC CTC CMC TGC CCA GAA TGC TGC CCT CCT GAT CCA GGA TCT CCA GCC     146
Ser Leu Xaa Cys Pro Glu Cys Cys Pro Pro Asp Pro Gly Ser Pro Ala
         -5                   1               5

TCC AGG GAC CCT AAC GTG GCC TGC GAA CGG                             176
Ser Arg Asp Pro Asn Val Ala Cys Glu Arg
 10                  15
```

(2) INFORMATION FOR SEQ ID NO: 62:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 287 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:

```
        (A) ORGANISM: Homo Sapiens
        (D) DEVELOPMENTAL STAGE: Fetal
        (F) TISSUE TYPE: brain

    (ix) FEATURE:
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: 36..101
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 6.7
                                seq LLLFQPSSHSATG/SS

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 62:

ATGAAACAAA ATATGTTCTA CCAGATAAAG TATTG ATG TCC CTA ACT GAT GTC     53
                                       Met Ser Leu Thr Asp Val
                                                   -20

CCC ATG TCC CTT CTG CTT TTT CAA CCC AGT TCC CAC AGT GCT ACT GGA   101
Pro Met Ser Leu Leu Leu Phe Gln Pro Ser Ser His Ser Ala Thr Gly
    -15                 -10                 -5

TCA TCT ATC AAA ATT ATA ATA CTT AAT TAC ATC ATA CTC CAG TTC AAA   149
Ser Ser Ile Lys Ile Ile Ile Leu Asn Tyr Ile Ile Leu Gln Phe Lys
  1               5                  10                  15

ACC CTT CAA ACA CTT CCT AAT GCT TTG AGG ATA CAC ATC AAA GTC TTT   197
Thr Leu Gln Thr Leu Pro Asn Ala Leu Arg Ile His Ile Lys Val Phe
            20                  25                  30

CAC ATT TAC TGT TCA TTT GTT TCC AGG TTT CAC TAT TAT AAA AAT ACT   245
His Ile Tyr Cys Ser Phe Val Ser Arg Phe His Tyr Tyr Lys Asn Thr
        35                  40                  45

GCC ACA GTA TTT TTC AGG TCA GTA TTA AAA AGG AGA ATG GGG           287
Ala Thr Val Phe Phe Arg Ser Val Leu Lys Arg Arg Met Gly
    50                  55                  60


(2) INFORMATION FOR SEQ ID NO: 63:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 453 base pairs
        (B) TYPE: NUCLEIC ACID
        (C) STRANDEDNESS: DOUBLE
        (D) TOPOLOGY: LINEAR

    (ii) MOLECULE TYPE: CDNA

    (vi) ORIGINAL SOURCE:
        (A) ORGANISM: Homo Sapiens
        (F) TISSUE TYPE: Substantia nigra

    (ix) FEATURE:
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: 238..288
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 6.6
                                seq LAFLLVSLYWSHM/HP
```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 63:

```
AATAGGGCTG AACCGAGGAC TGAAAAAGGG AGGAGGCAGA CCACTCGGAG AGGAGCTGGG   60

AAGCAGTGCA GAGAGGAGAG CGGASSNAGC TGCCGCTGAG CAAAGTGCGA CTGTATCTGG  120

TTTGGCCTGC TCTTCCTCAC CTTCCTCCTT TCCCTGAGCT GGCTGTACAT CGGGCTCGTC  180

CTTCTCAATG ACCTGCACAA CTTCAATGAA TTCCTCTTCC GCCGCTGGGG ACACTGG     237

ATG GAC TGG TCC CTG GCA TTC CTG CTG GTA TCT CTC TAC TGG TCA CAT    285
Met Asp Trp Ser Leu Ala Phe Leu Leu Val Ser Leu Tyr Trp Ser His
        -15                 -10                     -5

ATG CAT CCT TGC TAT TGG TCC TGG CCC TGC TCC TGC GGC TTT GTA GAC    333
Met His Pro Cys Tyr Trp Ser Trp Pro Cys Ser Cys Gly Phe Val Asp
      1               5                  10                  15

AGC CCC TGC ATC TGC ACA GCC TCC ACA AGG TGC TGC TGC TCC TCA TTA    381
Ser Pro Cys Ile Cys Thr Ala Ser Thr Arg Cys Cys Cys Ser Ser Leu
                 20                  25                  30

TGC TGC TTG TGG CGG CTG GCC TTG TGG GAC TGG ACA TCC AAT GGC AGC    429
Cys Cys Leu Trp Arg Leu Ala Leu Trp Asp Trp Thr Ser Asn Gly Ser
             35                  40                  45

AGG AGT GGC ATA GCT TGC GTG TGT                                    453
Arg Ser Gly Ile Ala Cys Val Cys
        50                  55
```

(2) INFORMATION FOR SEQ ID NO: 64:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 159 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 103..147
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.5
seq LLILFFMVGRIIP/SP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 64:

```
ATATTATATT TTTTAGTCAG ACTCACAGAA TTGAGTTGAT TTTATTCCTC ATTGGGTGGC   60

ACACTATTAT CGTTCTTCCC AAMCTTGTTC AGTATTTGTW TT ATG TAT TTA CTA    114
```

```
                                             Met Tyr Leu Leu
                                             -15

ATA CTT TTC TTT ATG GTA GGC AGA ATT ATC CCT TCC CCC CAC CGG       159
Ile Leu Phe Phe Met Val Gly Arg Ile Ile Pro Ser Pro His Arg
    -10                 -5                   1
```

(2) INFORMATION FOR SEQ ID NO: 65:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 245 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 78..221
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.5
seq LLVVSCCLLFHQA/IH

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 65:

```
AATGGATTAT GTGCTATGTG CACTCCACAG CCTAAGAACT ACAATCTTGG TTGTTGTTAA    60

TTTTTGATAC CTAAAAG ATG AAT AAG CCT CCC TGG GAA GAA TCC TGG GGG      110
                   Met Asn Lys Pro Pro Trp Glu Glu Ser Trp Gly
                               -45                 -40

CAA AAT CAA CTC TCA GGT GAG CCC GCT ACT TGG AGT CTT TGC ATT TCT     158
Gln Asn Gln Leu Ser Gly Glu Pro Ala Thr Trp Ser Leu Cys Ile Ser
        -35                 -30                 -25

CCT TTG CCT GGA AGA GAG CCG AGT TTA CTT GTT GTG TCT TGT TGC CTC     206
Pro Leu Pro Gly Arg Glu Pro Ser Leu Leu Val Val Ser Cys Cys Leu
    -20                 -15                 -10

TTG TTT CAT CAG GCT ATC CAT AAT AAA TTA CTG TGG CGG                 245
Leu Phe His Gln Ala Ile His Asn Lys Leu Leu Trp Arg
 -5                  1               5
```

(2) INFORMATION FOR SEQ ID NO: 66:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 299 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 177..248
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.5
seq FLILLSIDSLVSG/FL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 66:

```
AATTTTTTGC GACTTATTGA GGTTGTCTTT TCTGTCATTC AGGAAAGTTT TATGGTTTTC   60

TCTGGAAAGG TCTTGCTTGT TTCTTGTTAG GATTCTTGAT ACTTTTGAGT CTGTTTCCTC  120

CTTTGGTGTG TCATAATGGT TATTGATGGT GTATTGGGAA GTTATAGATA ATTTGG ATG  179
                                                              Met

TTG ATC TTA GAG CTA ACA ATG ATG CTG AGC TTT CTA ATT CTA TTG TCA    227
Leu Ile Leu Glu Leu Thr Met Met Leu Ser Phe Leu Ile Leu Leu Ser
            -20                 -15                 -10

ATT GAT TCT CTT GTA TCG GGT TTT TTA AGT AAG CGA AAA GGT CTG CGC    275
Ile Asp Ser Leu Val Ser Gly Phe Leu Ser Lys Arg Lys Gly Leu Arg
         -5                   1               5

GTC TGT GAT GGA AGC CGG TCC GGG                                    299
Val Cys Asp Gly Ser Arg Ser Gly
 10                  15
```

(2) INFORMATION FOR SEQ ID NO: 67:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 350 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 183..338
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.4
seq CLLGAAWASRLRT/QP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 67:

```
AATAGACGAG TAGACCTTAG GCGAGGGAAG GCATGTACTT TCCCTAAGAA GCGGGAGGAG   60

AGAAATAAAT AGAGGGCAGT GGGATGAGAA GAACCTCGCG GGATGGGAAG CCGCGGAGGG  120

AAGGGCCGTC TTTGGTTACC TGGAGAGCGG GAGCAGCTGC GGATCCCTTA TGAAGTGCCC  180

GG ATG AAG CTC CAG CGC TCA CGC GCT TTC CGC ATT GAG TGC AGC GCC      227
   Met Lys Leu Gln Arg Ser Arg Ala Phe Arg Ile Glu Cys Ser Ala
           -50                 -45                 -40

ATC TTG AGA AGG GCG GAG CGT CTT GTK TGG AAT GAC GTC TGT TCA GAG     275
Ile Leu Arg Arg Ala Glu Arg Leu Val Trp Asn Asp Val Cys Ser Glu
        -35                 -30                 -25

AGC CAA TCC CAG TCT CGC GAC TCC TGC TTG CTG GGC GCG GCT TGG GCC     323
Ser Gln Ser Gln Ser Arg Asp Ser Cys Leu Leu Gly Ala Ala Trp Ala
    -20                 -15                 -10

TCC AGG CTG CGC ACG CAG CCG CAT CCG                                 350
Ser Arg Leu Arg Thr Gln Pro His Pro
 -5                      1
```

(2) INFORMATION FOR SEQ ID NO: 68:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 295 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 236..283
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.7
seq FTLCVFTLPFLCA/CL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 68:

```
ACAAAAGGGA ATGTATTAGT GTCCCAGGAC TACTGTGACA AGGCACAAAA AACTGGGTGC   60

TTTACAACAA GAAAACTGTA CAGTGCTGGA GGCTAGAAGT CAAAATCAAG TGATTGGTTG  120

GACCATGCTC TCTCTGACAG TGACAGGGGA GAACCCTCCC TCGCCTCTCC TGGCTTCTGG  180

TATGCACCAG CAATTCCTGG CGTTCCTTGG CTCCTAGAAG CATCACTCCT ATCAC ATG   238
                                                             Met

GTC ATC TTC ACC CTG TGT GTC TTC ACA CTA CCC TTT CTC TGT GCA TGT    286
Val Ile Phe Thr Leu Cys Val Phe Thr Leu Pro Phe Leu Cys Ala Cys
-15                 -10                  -5                   1
```

```
CTG CCC AGG                                                          295
Leu Pro Arg
```

(2) INFORMATION FOR SEQ ID NO: 69:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 285 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 184..240
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.7
seq VLVVGTWSSQGQA/NS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 69:

```
AAAACTCTAG AAATCTGTGT TTCCGGAAAG CAGGGGTGGG AGCCGCCCAG TGCCTTCCCC    60

AAAACCCAAC ACTGAGATTT CAATGRTATG KTTTGKGAGT CTTTTTAAAA ATTCTCTCTG   120

GTTGCTGGGC TCAGTGACCA CGGTCAGGTT TGGAAGAGCA CCAGGCTGTG CGGGCCGAGG   180

CGG ATG TGG GGA GCA CTT CCT GTC CTC GTG GTG GGA ACC TGG TCC AGC     228
    Met Trp Gly Ala Leu Pro Val Leu Val Val Gly Thr Trp Ser Ser
                    -15                 -10                 -5

CAA GGG CAG GCG AAC AGC TGT GCG GGG CGG GGG ATG GGG CCA GAC GTG     276
Gln Gly Gln Ala Asn Ser Cys Ala Gly Arg Gly Met Gly Pro Asp Val
                  1               5                  10

TGT GGA GCG                                                         285
Cys Gly Ala
         15
```

(2) INFORMATION FOR SEQ ID NO: 70:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 247 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 143..190
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.7
seq LVCGFLQISLSLA/SL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 70:

```
AGTGACTGTA CACTGGGTTT CTTTGGAAAC TGCGTCTTTT CTCAATAATG GCAGGATCCC    60

CGATTTTAGA AAGTGGGCAG TGCTTTGTGT TACAGGGTTT GGGCTAAAGA CTGTTTGATA   120

ACCAATATTT TACAAGAATT AA ATG ACC AGA TTG GTC TGT GGC TTT CTC CAA   172
                         Met Thr Arg Leu Val Cys Gly Phe Leu Gln
                             -15                 -10

ATT TCC TTA TCC CTA GCT TCA TTG TTT CTG ACA ATT CCT CTA ATG TGG     220
Ile Ser Leu Ser Leu Ala Ser Leu Phe Leu Thr Ile Pro Leu Met Trp
     -5                  1               5                  10

TAC ATG CAA TCA AAA TGG TGG CGT GGG                                 247
Tyr Met Gln Ser Lys Trp Trp Arg Gly
                15
```

(2) INFORMATION FOR SEQ ID NO: 71:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 114 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 55..99
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.6
seq FLLPLLLHHLTFH/GR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 71:

```
AATCTATTTG TCCTCTAGAA ACTCTTCATA ATGCCCTATC TTTAAAGCAA GTGG ATG      57
                                                            Met
                                                            -15
```

```
AAT TTC TTG CTT CCA TTG CTT CTC CAT CAT CTG ACG TTC CAC GGA AGA      105
Asn Phe Leu Leu Pro Leu Leu Leu His His Leu Thr Phe His Gly Arg
                -10                 -5                   1

CCG CTG AAG                                                          114
Pro Leu Lys
          5
```

(2) INFORMATION FOR SEQ ID NO: 72:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 367 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 125..298
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.5
seq LIIFICXTASISA/YM

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 72:

```
CTGCTTCACT TTCAGGTTTC TCGAAGTGCC TTCTTGCTCC TGTCTGTTTC CCCATCCTGC     60

CAGATTTCTG TTTCTCTTGC TGGGCTTTTG GCAGTAGGGG GCTGTGTTGG TGGGCCCTAC    120

GAAG ATG CTC AGT GCT CGA GAT CGC CGG GAC CGG CAC CCT GAG GAG GGG     169
     Met Leu Ser Ala Arg Asp Arg Arg Asp Arg His Pro Glu Glu Gly
                 -55                 -50                 -45

GTA GTT GCA GAG CTC CAG GGC TTC GCG GTG GAC AAG GCC TTC CTC ACC      217
Val Val Ala Glu Leu Gln Gly Phe Ala Val Asp Lys Ala Phe Leu Thr
            -40                 -35                 -30

TCC CAC AAG GGC ATC CTG CTG GAA ACC GAG CTG GCC CTG ACC CTC ATC      265
Ser His Lys Gly Ile Leu Leu Glu Thr Glu Leu Ala Leu Thr Leu Ile
        -25                 -20                 -15

ATC TTC ATC TGC DTC ACG GCC TCC ATC TCT GCC TAC ATG GCC GCG GCG      313
Ile Phe Ile Cys Xaa Thr Ala Ser Ile Ser Ala Tyr Met Ala Ala Ala
    -10                  -5                   1               5

CTA CTG GAG TTC TTC ATC ACA CTT GCC TTC CTC TTC CTC TAT GCC ACC      361
Leu Leu Glu Phe Phe Ile Thr Leu Ala Phe Leu Phe Leu Tyr Ala Thr
                 10                  15                  20

CCA GCG                                                              367
Pro Ala
```

(2) INFORMATION FOR SEQ ID NO: 73:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 338 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 225..263
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.5
seq MLTMSVTLSPLRS/QD

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 73:

```
AAGATGTGCT GTAAACATCA AAAGAAGACG GTGGGATCAG GAGATGTTTG GACAGCTCTT    60

TATTCAGAAC ATCAAGGACT GTTGACAGTT GAATAAHNAG GGCGGAGGCT TATGGGATTG   120

CTAATGAGAT ACAAAGCCAC CTTGGAATAA AAATAAAHTT CTCTCTGTTG GCTCCTCCGG   180

CCATGGAGAG CTGTTTTVGA AAGAAGTGAG GTTTAGACTT CTCC ATG TTA ACC ATG   236
                                                 Met Leu Thr Met
                                                             -10

AGC GTG ACA CTT TCC CCC CTG AGG TCA CAG GAC CTG GAT CCC ATG GCT     284
Ser Val Thr Leu Ser Pro Leu Arg Ser Gln Asp Leu Asp Pro Met Ala
                 -5                   1                5

ACT GAT GCT TCA CCC ATG GCC ATC AAC ATG ACA CCC ACT GTG GAG CAG     332
Thr Asp Ala Ser Pro Met Ala Ile Asn Met Thr Pro Thr Val Glu Gln
         10                  15                  20

GGA CTG                                                             338
Gly Leu
     25
```

(2) INFORMATION FOR SEQ ID NO: 74:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 279 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 100..237
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.4
seq FFLLISSVRPISQ/TF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74:

```
ATTTAAATTT TGTGATCAAA GATGCCAAAT GGACACAACA CCATACCCTA GTTATTTCCC   60

TACTGTGTTT TCCTACAAGC TTACCTGCAT TTAGTAMCC ATG TTT ATR CCC GTA     114
                                           Met Phe Xaa Pro Val
                                               -45

GCA CTG ATC TTC CCC ATC TCA GTA AGT GAC CCT ACC ATT CAC CCT ATT    162
Ala Leu Ile Phe Pro Ile Ser Val Ser Asp Pro Thr Ile His Pro Ile
    -40                 -35                 -30

ACT CAA GCC CAG AAC CTA GAA AGC NTC CTA CAG TCC TTC TTT CTT CTA    210
Thr Gln Ala Gln Asn Leu Glu Ser Xaa Leu Gln Ser Phe Phe Leu Leu
-25                 -20                 -15                 -10

ATA TCA TCT GTA AGA CCC ATT AGT CAA ACC TTC AAA ATA GAT CTT TCT    258
Ile Ser Ser Val Arg Pro Ile Ser Gln Thr Phe Lys Ile Asp Leu Ser
                -5                   1               5

CCA TCT GTG CGG GCH ACC GGG                                        279
Pro Ser Val Arg Ala Thr Gly
        10
```

(2) INFORMATION FOR SEQ ID NO: 75:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 192 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 91..183
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.4
seq WCAVLRSWLAASS/AP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 75:

```
ATGCAAAAAA GCACTGTGCC TTCTGTCATT GAATCCTAGA ACACTGAGAT GAGAAAAACA   60

CTGTAGATGT ATAGGTTAGA TATATAGGGA ATG TTA TTA TTT TTT CCT TTT TTT   114
                                 Met Leu Leu Phe Phe Pro Phe Phe
                                     -30                 -25

GGA GAG ACG GTC TCG CTC CAT CAC CCA TGC TGG TGT GCA GTG CTG CGA   162
Gly Glu Thr Val Ser Leu His His Pro Cys Trp Cys Ala Val Leu Arg
            -20                 -15                 -10

TCG TGG CTC GCT GCA TCC TCC GCC CCT CGG                           192
Ser Trp Leu Ala Ala Ser Ser Ala Pro Arg
        -5                  1
```

(2) INFORMATION FOR SEQ ID NO: 76:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 199 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Surrenals

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 32..136
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.4
seq LVVVCYLSWRVSS/RS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 76:

```
ATACTCAAAA TCTACATTTC AGTTAGTTGA T ATG CCT CTT AAG AAT CTG TTC      52
                                   Met Pro Leu Lys Asn Leu Phe
                                   -35                 -30

TCT GTT GGT CTG TGG GAT CCT TAC AAT TTA CTG AAG AAA CAT GTT TTG   100
Ser Val Gly Leu Trp Asp Pro Tyr Asn Leu Leu Lys Lys His Val Leu
            -25                 -20                 -15

GTT GTT GTC TGC TAT TTA TCC TGG AGA GTG TCT TCC AGA AGT TGG ACT   148
Val Val Val Cys Tyr Leu Ser Trp Arg Val Ser Ser Arg Ser Trp Thr
        -10                 -5                  1

TTG CTG ATT ACA CCT GTA ACA CTT CAT GCT TCT CTG TCC ACC CAG GCC   196
Leu Leu Ile Thr Pro Val Thr Leu His Ala Ser Leu Ser Thr Gln Ala
  5                 10                  15                  20

CGG                                                               199
Arg
```

(2) INFORMATION FOR SEQ ID NO: 77:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 418 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 209..265
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.4
seq LSHLLPSLRQVIQ/EP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 77:

```
ATGTTTGATT GGTGTGCCTG AAAGTGAAGG GGAGAATGAA AACATCTGCC AGCTCCGCCG   60

ACCCGGCCCC CGCGGCCCTC CCAGCTCGGC TCCGGCTCAG TGGACAGGAA CCACTGAAGT  120

TTGCCTGACA CCATCAACCA GGCCCTAGTC ACCTGGCTTT GCCTTTGCCC TGCTGTGTGA  180

TCTTAGCTCC CTGCCCAGGC CCACAGCC ATG GCC ATG GCC CAG AAA CTC AGC     232
                               Met Ala Met Ala Gln Lys Leu Ser
                                               -15

CAC CTC CTG CCG AGT CTG CGG CAG GTC ATC CAG GAG CCT CAG CTA TCT    280
His Leu Leu Pro Ser Leu Arg Gln Val Ile Gln Glu Pro Gln Leu Ser
    -10                 -5                   1                5

CTG CAG CCA GAG MYG GTC TTC ACG GTG GAT CGA GCT GAG GTG CCG CCG    328
Leu Gln Pro Glu Xaa Val Phe Thr Val Asp Arg Ala Glu Val Pro Pro
                10                  15                  20

CTC TTC TGG AAG CCG TAC ATC TAT GCG GGC TAM CGG CCG CTG CAT CAG    376
Leu Phe Trp Lys Pro Tyr Ile Tyr Ala Gly Xaa Arg Pro Leu His Gln
            25                  30                  35

ACC TGG CGC TTC TAT TTC CGC ACG CTG TTC CAG CAG CAC AAC            418
Thr Trp Arg Phe Tyr Phe Arg Thr Leu Phe Gln Gln His Asn
        40                  45                  50
```

(2) INFORMATION FOR SEQ ID NO: 78:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 370 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 149..232
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.4
seq LALVALAPHSVQK/SX

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 78:

```
ATGCCTCGAC CCCTTTAGAA GTTCCTTTTC CCCAGTCATC CTCTGGAACA GCCCTACCTT    60

TGGGAACTGC CCCTGAAGCC CCAAHCCTTC CTAMCCARAC CTAATAGGGC CTCCCATCTC   120

CCCAGCTGCC TTAGCTCTAG CCTCTCCC ATG ATA GCT CCA ACT CTG AAA GGG      172
                               Met Ile Ala Pro Thr Leu Lys Gly
                                           -25

ACC CCT TCC TCT TCA GCT CCC TTA GCT CTG GTT GCC CTG GCT CCC CAC     220
Thr Pro Ser Ser Ser Ala Pro Leu Ala Leu Val Ala Leu Ala Pro His
-20                 -15                 -10                  -5

TCA GTT CAG AAG AGT TYN VCT TTT CCA CCT AAC CTT CTT ACT TCA CCT     268
Ser Val Gln Lys Ser Xaa Xaa Phe Pro Pro Asn Leu Leu Thr Ser Pro
                 1               5                  10

CCT TCA GTG GCT GHA GCT GAG TCA GGG TCA GTG ATA ACT CTG TCA GCT     316
Pro Ser Val Ala Xaa Ala Glu Ser Gly Ser Val Ile Thr Leu Ser Ala
        15                  20                  25

SCC ATT GCT CCC TCA GAA CCA AAG ACT AAT CTT AAT AAA GTT CCC TCT     364
Xaa Ile Ala Pro Ser Glu Pro Lys Thr Asn Leu Asn Lys Val Pro Ser
    30                  35                  40

GAG GTA                                                             370
Glu Val
 45
```

(2) INFORMATION FOR SEQ ID NO: 79:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 208 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens

(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 86..196
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.3
seq LVESLCLVFNLLS/LP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 79:

```
AGATATGTCA ACACCTAAGA ATCAGAACTA GTATCTTGAA AAGTTAGGAG CCCTGGGGTT    60

TTGTTTTTGC TTTTGCTTTT AAAGG ATG TGT TTG TTC CCT GTA TCA CCG TGC   112
                            Met Cys Leu Phe Pro Val Ser Pro Cys
                                    -35                 -30

CCA GCT TAC TCC TTT TCT TCG GAA ASR STT GGT GCC GTA TTG TTA CTG   160
Pro Ala Tyr Ser Phe Ser Ser Glu Xaa Xaa Gly Ala Val Leu Leu Leu
        -25                 -20                 -15

GTT GAA TCT CTG TGT TTG GTT TTT AAT TTG TTA TCT TTG CCC CCA AGG   208
Val Glu Ser Leu Cys Leu Val Phe Asn Leu Leu Ser Leu Pro Pro Arg
    -10                 -5                  1
```

(2) INFORMATION FOR SEQ ID NO: 80:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 412 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 140..184
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.3
seq IAVLFCFLLLIIF/QT

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 80:

```
AAACATTCCC TTCTGTCCTT TCTTTGTTTT TAAAGAAAGC TCTGATTTTG TTTCATTTTC    60

AGCTGGAGAC TTAAATGACA CCAAGCAAAG CCTACTTAGT TTAGATCTCC AGAAATTGGC   120

TGGTGGAAVN RAATCAAAC ATG AAG ATT GCA GTT TTG TTT TGT TTT TTG CTG   172
                     Met Lys Ile Ala Val Leu Phe Cys Phe Leu Leu
                     -15                 -10                 -5
```

```
CTT ATC ATT TTT CAA ACT GAC TTT GGA AAA AAT GAA GAA ATT CCT AGG    220
Leu Ile Ile Phe Gln Thr Asp Phe Gly Lys Asn Glu Glu Ile Pro Arg
                1               5                   10

AAG CAA AGG AGG AAG ATC TAC CAC AGA AGG TTG AGG AAA AGT TCA ACV    268
Lys Gln Arg Arg Lys Ile Tyr His Arg Arg Leu Arg Lys Ser Ser Thr
        15                  20                  25

TCA CAC AAG CAC AGA TCA AAC AGA CAG CTT GGA ATT CCG CAA ACA ACA    316
Ser His Lys His Arg Ser Asn Arg Gln Leu Gly Ile Pro Gln Thr Thr
    30                  35                  40

GTT TTT ACA CCA GTA GCA AGA CTT CCT ATT GTT AAC TTT GAT TAT AGC    364
Val Phe Thr Pro Val Ala Arg Leu Pro Ile Val Asn Phe Asp Tyr Ser
45                  50                  55                  60

ATG GAG GAA AAG TTT GAA TCC TTT TCA AGT TTT CCT GGA GTA GAA TCA    412
Met Glu Glu Lys Phe Glu Ser Phe Ser Ser Phe Pro Gly Val Glu Ser
                65                  70                  75
```

(2) INFORMATION FOR SEQ ID NO: 81:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 169 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 80..160
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.2
seq VGAVLLSSLPISP/QY

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 81:

```
ACCCTACCAT CCATACCCTC CATTCTTACT CCGTGTCATG GAGAAACCTG CATAGTGCTG   60

GGGATTAGTG CCACAATTA ATG TGT AGT CCG AGG TCT CCC TTA AAT CTG TCT   112
                     Met Cys Ser Pro Arg Ser Pro Leu Asn Leu Ser
                                 -25                 -20

TTG GTC CCT GTC GGA GCA GTT CTG CTT AGC TCC CTC CCC ATT TCT CCA    160
Leu Val Pro Val Gly Ala Val Leu Leu Ser Ser Leu Pro Ile Ser Pro
    -15                 -10                 -5

CAG TAC GGG                                                        169
Gln Tyr Gly
  1
```

(2) INFORMATION FOR SEQ ID NO: 82:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 230 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 24..167
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.2
seq EVVTLPLTSHCLA/QV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 82:

```
ATGAATTTCT AAAACTCTTT TTT ATG GGA CTT CAC ATT TCT CTG ATT AAA TTT    53
                          Met Gly Leu His Ile Ser Leu Ile Lys Phe
                                      -45                    -40

CTT CTT GCT AAT GGA CCC CAT ATT CCT AGT CAC CAA AGA CCT TTT GAA     101
Leu Leu Ala Asn Gly Pro His Ile Pro Ser His Gln Arg Pro Phe Glu
            -35                 -30                 -25

CCT AAA GGG GAA AAA AGC TGC AGA ATT GAA GTG GTG ACT CTG CCA CTT     149
Pro Lys Gly Glu Lys Ser Cys Arg Ile Glu Val Val Thr Leu Pro Leu
        -20                 -15                 -10

ACT AGC CAT TGT CTT GCC CAA GTT GCA AGT TCT GAC CTC ATC CAT AGG     197
Thr Ser His Cys Leu Ala Gln Val Ala Ser Ser Asp Leu Ile His Arg
     -5                  1               5                  10

ATG AGA ACC ATA ACA GGT ACC TCG TCA CAC GGG                         230
Met Arg Thr Ile Thr Gly Thr Ser Ser His Gly
                 15                  20
```

(2) INFORMATION FOR SEQ ID NO: 83:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 349 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens

(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 158..235
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.1
seq LLTLYVFVASSMQ/IY

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 83:

```
ATACATGGAG GCATGTAATG AATACTGAAT GAATAAATTA TTAAAACCTC AAGCTTATTC   60

TAGGCTTAAG CTCTGGTTCT CAATAAAATA CTACCAAATA AACATAAAAA TACGTACTTA  120

CTTTAGAGCA TTTTGAAAGT ACATAACTTA AGGAAAA ATG AAA ACT ACC TAT GTA  175
                                         Met Lys Thr Thr Tyr Val
                                             -25

ATA TTT ATG CAA AGC AAA GCA CTA TTA ACA TTG TAT GTA TTT GTA GCC     223
Ile Phe Met Gln Ser Lys Ala Leu Leu Thr Leu Tyr Val Phe Val Ala
-20                 -15                 -10                  -5

TCT TCT ATG CAA ATT TAT GTA TTA CAC ATT TCA AAT TAC CCA ACA GAT     271
Ser Ser Met Gln Ile Tyr Val Leu His Ile Ser Asn Tyr Pro Thr Asp
                 1               5                  10

GAG CAT TTT CCT ATC ATT AAG CAT TTT TAT TTT ACT TTT AAA ATC CAC     319
Glu His Phe Pro Ile Ile Lys His Phe Tyr Phe Thr Phe Lys Ile His
         15                  20                  25

TTT AGT AAA ATT ATT TAT GTG CAG TAC AGT                             349
Phe Ser Lys Ile Ile Tyr Val Gln Tyr Ser
     30                  35
```

(2) INFORMATION FOR SEQ ID NO: 84:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 142 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 68..112
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.1
seq ALVFLIFLRFINI/SE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 84:

```
AATTACTCCC TTCACTGGAA AGAATAAAAT ATTACCACAT AAATTCGTAT TATGATAATG   60

GCCCACT ATG AAT GCT TTG GTG TTC TTA ATT TTC TTA AGA TTT ATT AAT    109
        Met Asn Ala Leu Val Phe Leu Ile Phe Leu Arg Phe Ile Asn
        -15                 -10                 -5

ATT TCT GAA GTA ACT ACT AAA TGC CAA GCA GGG                        142
Ile Ser Glu Val Thr Thr Lys Cys Gln Ala Gly
    1               5                   10
```

(2) INFORMATION FOR SEQ ID NO: 85:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 190 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 86..172
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.1
seq WGFLLTGHSLSHS/SK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 85:

```
AATTCCATCT CAGTTGCTGG TTTCACAGGC AGGACCACCC CTGGGCGCCT CTGTCCCCCG   60

GTCGGGGAGT CTGATCCTGC CTCCC ATG CAG CTG GGT CCC CTT CAC ACT GTG    112
                            Met Gln Leu Gly Pro Leu His Thr Val
                                            -25

TCT ACA CCC TTT TTC TTC TGT TGG GGT TTT CTC CTC ACT GGC CAT TCT    160
Ser Thr Pro Phe Phe Phe Cys Trp Gly Phe Leu Leu Thr Gly His Ser
-20                 -15                 -10                 -5

CTT TCY CAC TCC TCT AAG TCC TGC CAC CTG                            190
Leu Ser His Ser Ser Lys Ser Cys His Leu
                1                   5
```

(2) INFORMATION FOR SEQ ID NO: 86:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 226 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 41..211
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.1
seq VGSVCCCVGPLRG/LV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 86:

```
CTCTAGGACT GGCTTGGTGG AGGTBRATGA AGGCCGTGAG ATG GGG CGG GGG TGG      55
                                            Met Gly Arg Gly Trp
                                                        -55

GAG AGG ACA GTT TGC AGC CTT GGG TGG CGC GGG GGC CCA GAC CCT CTG     103
Glu Arg Thr Val Cys Ser Leu Gly Trp Arg Gly Gly Pro Asp Pro Leu
        -50                 -45                 -40

TCT TGG GCC ACT TGT TGG TCG GGA GCC AGA TCT CGC CAC ACC CGC GTT     151
Ser Trp Ala Thr Cys Trp Ser Gly Ala Arg Ser Arg His Thr Arg Val
    -35                 -30                 -25

TCC TCG ATT GTA AAC GGT TAT GTT GGG AGT GTT TGC TGC TGC GTG GGT     199
Ser Ser Ile Val Asn Gly Tyr Val Gly Ser Val Cys Cys Cys Val Gly
-20                 -15                 -10                  -5

CCG CTG AGA GGA TTA GTC TGG GCC CCT                                 226
Pro Leu Arg Gly Leu Val Trp Ala Pro
                1               5
```

(2) INFORMATION FOR SEQ ID NO: 87:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 195 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 115..180
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5
seq HLFVTWSSQRALS/HP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 87:

```
AACCTGCCAG TKATGCAAAT GCCAAAATGT GGGTCATCAT ATAGTATATT TGAAACCTTT   60

CTGAACATGT ACACCACCCA ATGCTAGAGG CTGACTTGGA AACCGGTGGG TGCA ATG    117
                                                            Met

CCC GAG GCT GTG GAA CAA TCA GCC CAT CTC TTT GTG ACC TGG AGC AGT    165
Pro Glu Ala Val Glu Gln Ser Ala His Leu Phe Val Thr Trp Ser Ser
    -20                 -15                 -10

CAG AGG GCC CTC AGT CAC CCC GCC CCA TTG                            195
Gln Arg Ala Leu Ser His Pro Ala Pro Leu
 -5                  1               5
```

(2) INFORMATION FOR SEQ ID NO: 88:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 386 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 240..311
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5
seq CWLIALSVPLVFW/VT

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 88:

```
AATTTGTTTA GAAACATGCC TTTCATTAAT GTAATTCCCT GCATTTACAG ATAGACAATC   60

TGAGCCAGGA CTAGCAGTCC AGGATTCCKG ATGTCTTACC TTGTCCTTAA CTCTGATACC  120

ATCCTACTGA CCAAAGTTNG CACTTCTTTG GACTTTTTTA GAGACTACCA AATGGGCTTT  180

GACACGTTCA TTTGGAATAC ATCAAGCCTT AATATTTATA TCTTATATCA TTGAACATA   239

ATG CCT GGC ACA CAC ACG TTT ACT TTT AAA AGC TGC TGG CTT ATA GCT    287
Met Pro Gly Thr His Thr Phe Thr Phe Lys Ser Cys Trp Leu Ile Ala
            -20                 -15                 -10

CTT TCT GTG CCT TTG GTC TTT TGG GTG ACA TTC TGG CCG TGT AAC TTT    335
Leu Ser Val Pro Leu Val Phe Trp Val Thr Phe Trp Pro Cys Asn Phe
             -5                   1               5

TAT CCG TCC CTT GAC TTC TGC ATG TTA ACT AAG RCT AAG AGC ATA TTC    383
Tyr Pro Ser Leu Asp Phe Cys Met Leu Thr Lys Xaa Lys Ser Ile Phe
```

```
    10                  15                  20

ATA                                                                386
Ile
 25
```

(2) INFORMATION FOR SEQ ID NO: 89:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 197 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 69..134
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.9
seq LFCLIGLDLLCQV/FS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 89:

```
ATAAATCTCA CAATTAGAAT TGCATAATTG ATTCACATCA ATAATAGACC TATATTTGAT    60

GTTCAGCT ATG CTT TTA CTG ACT TTC AAA TGG TTT CTG TTT TGC TTA ATT    110
         Met Leu Leu Leu Thr Phe Lys Trp Phe Leu Phe Cys Leu Ile
                 -20                 -15                 -10

GGT TTG GAT TTG CTC TGT CAG GTT TTT TCC CCT TAT TTC CTT AGT GAG    158
Gly Leu Asp Leu Leu Cys Gln Val Phe Ser Pro Tyr Phe Leu Ser Glu
             -5                   1              5

AAA GTT GCA GAC CTG CTG TTT TAT ATG TCT CTG TTT TTT                197
Lys Val Ala Asp Leu Leu Phe Tyr Met Ser Leu Phe Phe
     10                  15                  20
```

(2) INFORMATION FOR SEQ ID NO: 90:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 169 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal

```
        (F) TISSUE TYPE: brain

    (ix) FEATURE:
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: 5..160
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 4.9
                                seq VPNLHLLLPLTTP/QP

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 90:

ATTA ATG GCA ACY ACA GGC AGA CGG CAG GCT GAA CCT CCG CCC GTC CGG      49
     Met Ala Thr Thr Gly Arg Arg Gln Ala Glu Pro Pro Pro Val Arg
             -50                -45                 -40

CCC GCC CAT TCC CGA CCT CCA CCT AGG GTG CCT GGG AGC AGC AGT CTA       97
Pro Ala His Ser Arg Pro Pro Pro Arg Val Pro Gly Ser Ser Ser Leu
        -35                 -30                 -25

GGG CTG GCA GGA CTT ATG TCC CCC GTC CCC AAC CTT CAC CTA CTC CTC      145
Gly Leu Ala Gly Leu Met Ser Pro Val Pro Asn Leu His Leu Leu Leu
    -20                 -15                 -10

CCC CTT ACT ACT CCC CAA CCT CGG                                      169
Pro Leu Thr Thr Pro Gln Pro Arg
 -5                  1


(2) INFORMATION FOR SEQ ID NO: 91:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 185 base pairs
        (B) TYPE: NUCLEIC ACID
        (C) STRANDEDNESS: DOUBLE
        (D) TOPOLOGY: LINEAR

    (ii) MOLECULE TYPE: CDNA

    (vi) ORIGINAL SOURCE:
        (A) ORGANISM: Homo Sapiens
        (F) TISSUE TYPE: Substantia nigra

    (ix) FEATURE:
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: 27..74
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 4.9
                                seq FIYLQAHFTLCSG/WS

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 91:

AATAGATAGG AGAGCAAGCC TCACCA ATG GTT CCC TTC ATC TAT CTG CAA GCC      53
                             Met Val Pro Phe Ile Tyr Leu Gln Ala
                                 -15                 -10

CAC TTT ACA CTC TGT TCT GGG TGG TCC AGC ACA TAC CGG GAC CTC CGG      101
His Phe Thr Leu Cys Ser Gly Trp Ser Ser Thr Tyr Arg Asp Leu Arg
         -5                  1                   5
```

```
AAG GGT GTG TAT GTG CCC TAC ACC CAG GGC AAG TGG GAA GGG GAG CTG      149
Lys Gly Val Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu
 10                 15                  20                  25

GGC ACC GAC CTG GTA AGC ATC CCC CAT GGC CCA AAG                      185
Gly Thr Asp Leu Val Ser Ile Pro His Gly Pro Lys
                30                  35
```

(2) INFORMATION FOR SEQ ID NO: 92:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 239 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 150..191
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.8
seq FFSFLLTINLVSL/QV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 92:

```
ATAACATTAC AGTTTGGGCT CTTGGTGCCC AATGATTGAT TTATCTATAG TATAGATTTA    60

TTTCTCACAG TACCTCTTGG AATGCTCATT TTTAACCCCA ATAGTTAAAT TTGCCTTGGT   120

AAGCTACAAA AACAGGCACC AAAGCAGCA ATG TTT TTT AGT TTT CTG TTG ACC      173
                                Met Phe Phe Ser Phe Leu Leu Thr
                                                -10

ATA AAT CTC GTT TCT TTA CAA GTA GTA ATT CTA AAC AGA GTA TAC CTT      221
Ile Asn Leu Val Ser Leu Gln Val Val Ile Leu Asn Arg Val Tyr Leu
     -5                  1                   5                  10

AAC CAG CCA GAT GCA CGG                                              239
Asn Gln Pro Asp Ala Arg
                 15
```

(2) INFORMATION FOR SEQ ID NO: 93:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 168 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 100..162
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.8
seq NWGLLCFASECTT/DR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 93:

```
AGAGGATTGG AGAGGAAGAG TTGGATTCCA GAAATATGCA GGAGGTGGCC GCAGGAATTG    60

GGAGATAAGG CAGAGGGAAG AGGCCATGCC CAGTTTCTG ATG TGG CCT GGG CGT      114
                                          Met Trp Pro Gly Arg
                                              -20

GAA TGC AAA AAT TGG GGT CTT CTG TGC TTT GCA AGT GAG TGC ACC ACC     162
Glu Cys Lys Asn Trp Gly Leu Leu Cys Phe Ala Ser Glu Cys Thr Thr
    -15                 -10                     -5

GAT CGG                                                             168
Asp Arg
  1
```

(2) INFORMATION FOR SEQ ID NO: 94:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 238 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 113..223
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.8
seq CLLLTLRQPPTHS/GI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 94:

```
AAGTTTATTT ATCAAGTGAA ATAGGGTTAG TACTACCAGT CTCATAAGGT TATTATGAGA    60

TACCAAATGG AAATCTAAAG AAGATCYTTG CACCATAGTA TGTGATCAGT GA ATG TTA   118
                                                          Met Leu
```

```
ACT CCT TTC TCC TTG GAG GAG AAA CTT CTA GAA TGT CAT TAT GTA CTT      166
Thr Pro Phe Ser Leu Glu Glu Lys Leu Leu Glu Cys His Tyr Val Leu
-35                 -30                 -25                 -20

GCA AAA CTA GCT GGG GCA TGT CTC CTG TTG ACT CTA AGG CAG CCA CCC      214
Ala Lys Leu Ala Gly Ala Cys Leu Leu Leu Thr Leu Arg Gln Pro Pro
                -15                 -10                  -5

ACA CAT TCT GGA ATC CCA CAC GGG                                      238
Thr His Ser Gly Ile Pro His Gly
              1               5
```

(2) INFORMATION FOR SEQ ID NO: 95:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 204 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 79..198
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.7
seq IVFVGLIFLKSSA/HR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 95:

```
ATGATTGMGG GATTGTGTAG AGGTGATTTT GAAGATGGAA GACTTGTGCA CTGAAGAAAA     60

TGAGAAAAAT GAGAAGAA ATG AAA AGA ATA AAA TCA ATG ATG GGA AAA GTT      111
                    Met Lys Arg Ile Lys Ser Met Met Gly Lys Val
                    -40                 -35                 -30

GAA CAT ATA AAG ATT AAA GGA GAA AAA CAA AGA AGC CGT CAT GTA AAA      159
Glu His Ile Lys Ile Lys Gly Glu Lys Gln Arg Ser Arg His Val Lys
                -25                 -20                 -15

ATA GTA TTT GTT GGG CTT ATT TTT CTA AAA AGC AGT GCA CAT AGG          204
Ile Val Phe Val Gly Leu Ile Phe Leu Lys Ser Ser Ala His Arg
            -10                  -5                   1
```

(2) INFORMATION FOR SEQ ID NO: 96:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 136 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE

```
        (D) TOPOLOGY: LINEAR

    (ii) MOLECULE TYPE: CDNA

    (vi) ORIGINAL SOURCE:
        (A) ORGANISM: Homo Sapiens
        (F) TISSUE TYPE: Substantia nigra

    (ix) FEATURE:
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: 29..79
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 4.7
                                seq LCLFCKICPFTHG/VA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 96:

AACCAGAAAC TACTTTGCAT CCCATTGA ATG CAA TCT GCT CTG TGC CTC TTT      52
                               Met Gln Ser Ala Leu Cys Leu Phe
                                       -15                 -10

TGT AAA ATC TGC CCA TTT ACA CAT GGT GTT GCC ACC CCA GCC TGG GAA    100
Cys Lys Ile Cys Pro Phe Thr His Gly Val Ala Thr Pro Ala Trp Glu
                -5                   1                   5

CTG AGC AGC AAG AGG AAA GCT TCC CAC CCG CCC CGG                    136
Leu Ser Ser Lys Arg Lys Ala Ser His Pro Pro Arg
        10                  15


(2) INFORMATION FOR SEQ ID NO: 97:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 387 base pairs
        (B) TYPE: NUCLEIC ACID
        (C) STRANDEDNESS: DOUBLE
        (D) TOPOLOGY: LINEAR

    (ii) MOLECULE TYPE: CDNA

    (vi) ORIGINAL SOURCE:
        (A) ORGANISM: Homo Sapiens
        (F) TISSUE TYPE: Surrenals

    (ix) FEATURE:
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: 112..318
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 4.7
                                seq SLLCLAFLLGRFL/HM

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 97:

AGCTAGTAGC AGCTCTGGCA GAAGCAACGG TGGCTTCGAG GGATGGCGGC GGCTGCAACA    60

GGACCTGCAG CATCCCAGAG GAACTGACTA AGACTTTGGA ACAGAAACCA G ATG ATG   117
                                                         Met Met
```

```
CAC AAT ATT ATC GTC AAA GAG CTT ATT GTC ACA TTC TTC TTG GGA ATT      165
His Asn Ile Ile Val Lys Glu Leu Ile Val Thr Phe Phe Leu Gly Ile
        -65                 -60                 -55

ACT GTG GTG CAG ATG CTA ATT TCA GTG ACT GGA TTA AAA GGT GTC GAA      213
Thr Val Val Gln Met Leu Ile Ser Val Thr Gly Leu Lys Gly Val Glu
    -50                 -45                 -40

GCT CAG AAT GGC TCG GAA TCT GAG GTG TTT GTG GGG AAG TAT GAG ACC      261
Ala Gln Asn Gly Ser Glu Ser Glu Val Phe Val Gly Lys Tyr Glu Thr
-35                 -30                 -25                 -20

CTC GTG TTT TAC TGG CCC TCG CTG CTG TGC CTT GCC TTC CTG CTG GGC      309
Leu Val Phe Tyr Trp Pro Ser Leu Leu Cys Leu Ala Phe Leu Leu Gly
                -15                 -10                 -5

CGC TTC CTG CAT ATG TTT GTC AAG GCT CTG AGG GTG CAC CTC GGC TGG      357
Arg Phe Leu His Met Phe Val Lys Ala Leu Arg Val His Leu Gly Trp
             1               5                  10

GAG CTC CAG GTG GAA GAA AAA TCT GTC CTG                              387
Glu Leu Gln Val Glu Glu Lys Ser Val Leu
    15                  20
```

(2) INFORMATION FOR SEQ ID NO: 98:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 324 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 19..99
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.6
seq RLLCSRLCQQLRS/KR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 98:

```
ATCATCTGAA TCTTAATC ATG TCC AAC TGC CTG CAA AAT TTC CTG AAA ATT      51
                    Met Ser Asn Cys Leu Gln Asn Phe Leu Lys Ile
                            -25                 -20

ACA AGC ACT CGT CTT CTA TGT TCA AGA TTA TGC CAA CAG TTA AGA AGT      99
Thr Ser Thr Arg Leu Leu Cys Ser Arg Leu Cys Gln Gln Leu Arg Ser
    -15                 -10                 -5

AAA AGG AAG TTT TTC GGA ACT GTG CCA ATA TCC AGA TTG CAT AGG CGA      147
Lys Arg Lys Phe Phe Gly Thr Val Pro Ile Ser Arg Leu His Arg Arg
 1               5                  10                  15
```

```
GTT GTC ATT ACA GGC ATT GGC TTA GTG ACT CCT CTT GGT GTT GGA ACT      195
Val Val Ile Thr Gly Ile Gly Leu Val Thr Pro Leu Gly Val Gly Thr
            20                  25                  30

CAC CTG GTT TGG GAT CGT CTT ATC GGA GGA GAG AGT GGA ATT GTT TCA      243
His Leu Val Trp Asp Arg Leu Ile Gly Gly Glu Ser Gly Ile Val Ser
        35                  40                  45

CTG GTT GGT GAA GAG TAT AAG AGT ATC CCT TGC AGT GTT GCT GCT TAT      291
Leu Val Gly Glu Glu Tyr Lys Ser Ile Pro Cys Ser Val Ala Ala Tyr
    50                  55                  60

GTG CCA AGA GGT AGT GAT GAA GGT CAG TCA GGG                          324
Val Pro Arg Gly Ser Asp Glu Gly Gln Ser Gly
 65                 70                  75
```

(2) INFORMATION FOR SEQ ID NO: 99:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 241 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 74..121
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.6
seq XGLFLRTTAAARA/CR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 99:

```
AATCAGACGT GTGTGTGTCC CTGCGGCGCT AAGAAGGGGA GACTGAGGCT GMGGCTGGGG      60

AACATCGGGC AGC ATG AGC GGC NCM GGG CTC TTC CTG CGC ACC ACG GCT       109
               Met Ser Gly Xaa Gly Leu Phe Leu Arg Thr Thr Ala
                   -15                 -10                  -5

GCG GCT CGT GCC TGC CGG GGT CTG GTG GTC TCT ACC GCG AAC CGG CGG      157
Ala Ala Arg Ala Cys Arg Gly Leu Val Val Ser Thr Ala Asn Arg Arg
                    1               5                   10

CTA CTG CGC ACC AGC CCG CCT GTA CGA GCT TTC GCC AAA GAG CTT TTC      205
Leu Leu Arg Thr Ser Pro Pro Val Arg Ala Phe Ala Lys Glu Leu Phe
        15                  20                  25

CTA GGC AAA ATC RAG AAG GTA ACG CGA GCC CTG GGC                      241
Leu Gly Lys Ile Xaa Lys Val Thr Arg Ala Leu Gly
    30                  35                  40
```

(2) INFORMATION FOR SEQ ID NO: 100:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 207 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 112..186
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.5
seq LTWLHLLLSHLKS/SL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 100:

```
ACCTTCCCAC CACCCGCAGG AGCACCTGCC TCCATTCACA ACCGCCGACA CTACTGCAGC    60

CTCATCCTAA CTGACCTTGG CTTCTCTGGG AGACCCTTCC TGCACCCTGA A ATG AAT    117
                                                         Met Asn
                                                         -25

CCT CTG CAC AAG CAC TGT GCT GCA GGA CCC CTC ACC TGG CTC CAC CTC     165
Pro Leu His Lys His Cys Ala Ala Gly Pro Leu Thr Trp Leu His Leu
            -20                 -15                 -10

CTA CTT TCT CAC TTG AAG TCT TCT CTT GTA ATG CCT CCA AAG             207
Leu Leu Ser His Leu Lys Ser Ser Leu Val Met Pro Pro Lys
        -5                  1                   5
```

(2) INFORMATION FOR SEQ ID NO: 101:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 143 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 18..104

```
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 4.5
                                seq FAVLRVLHLPALT/AP

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 101:

AACACATCTT CCCTGAG ATG CCT AAG GAC AAA AGA GGA GCT AGA CAC AAC      50
                   Met Pro Lys Asp Lys Arg Gly Ala Arg His Asn
                                   -25                 -20

TCT CCC CAT TTT TCC TTT GCT GTC TTA AGA GTG CTC CAT CTT CCA GCA     98
Ser Pro His Phe Ser Phe Ala Val Leu Arg Val Leu His Leu Pro Ala
        -15                 -10                 -5

CTG ACT GCC CCT CTG TGG CTG GCT CCT TTC TCT ACC CTC CCC AGG        143
Leu Thr Ala Pro Leu Trp Leu Ala Pro Phe Ser Thr Leu Pro Arg
         1               5                  10


(2) INFORMATION FOR SEQ ID NO: 102:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 269 base pairs
        (B) TYPE: NUCLEIC ACID
        (C) STRANDEDNESS: DOUBLE
        (D) TOPOLOGY: LINEAR

    (ii) MOLECULE TYPE: CDNA

    (vi) ORIGINAL SOURCE:
        (A) ORGANISM: Homo Sapiens
        (F) TISSUE TYPE: Substantia nigra

    (ix) FEATURE:
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: 126..242
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 4.5
                                seq LCVSRQLLTGART/LF

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 102:

AAGAACAAGC AGCTGACATG ATTGCTGTCG GTGTTCACAG AGCCAAATAT GAAGATAATT   60

GTAAAGAAAG AGACGGTCGC ATCGGATAGA AGATGTGATC CTGCTCTCAC GTTTTTCCTT  120

CTGGC ATG ACC ATA CAC GTT TTG AGA AAA TGT TGC CAA ATG GGT AGA CTA  170
      Met Thr Ile His Val Leu Arg Lys Cys Cys Gln Met Gly Arg Leu
                  -35                 -30                 -25

AAC AAT GAA TGG CTG CCG GGT TTA GTC ATA CCT CTC TGT GTG AGC CGT    218
Asn Asn Glu Trp Leu Pro Gly Leu Val Ile Pro Leu Cys Val Ser Arg
                -20                 -15                 -10

CAA TTG CTG ACG GGA GCT AGG ACA TTA TTC CAG CTA CAA AAT GGG CCC    266
Gln Leu Leu Thr Gly Ala Arg Thr Leu Phe Gln Leu Gln Asn Gly Pro
            -5                   1               5
```

```
GCG                                                                  269
Ala
```

(2) INFORMATION FOR SEQ ID NO: 103:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 348 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 253..342
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.5
seq IAALLGLLQLRFK/AE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 103:

```
GTCCAGATTG GGCCACTTCT TTCTCAGCTC TAATGACTTT CCTCAGTTCC GTGGGTTACT   60

CCTGCCAACT CGACGCCGGC CGCCATGACA CTCGCTCGGA AAGCGGCAGC GNATCATAGA  120

AAAGCGCCGC GGTGGCGTAG ACAGGCCCCG CGAASCGCCG GACGTGTCCT TGGCGCAAGG  180

GAGGCTGGGA TCGCGGAGGA CCGAGCGCGG GCTGGATTAA CCGCAGCCAG TGCCTAGCGC  240

AAGGTTAGGT GC ATG CAG GCG GCC AGC TTC GGC CGG GGA AGG AAT GGC CTG  291
              Met Gln Ala Ala Ser Phe Gly Arg Gly Arg Asn Gly Leu
              -30                 -25                 -20

GAT AAC TGG GGC ATC GCG GCG CTC CTC GGC CTC CTG CAG CTG CGT TTC    339
Asp Asn Trp Gly Ile Ala Ala Leu Leu Gly Leu Leu Gln Leu Arg Phe
    -15                 -10                      -5

AAA GCA GAG                                                        348
Lys Ala Glu
    1
```

(2) INFORMATION FOR SEQ ID NO: 104:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 465 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 310..438
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.5
seq ENLLLCCHRCTNC/QR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 104:

```
AAGGTTCTTT CTCATTTGTG TAGGTGGATG TTCCTTCAAT CTTTGAAGTT GCTGTCCCTT   60

GGATGGATTT TTGTTGCTTT CATCTTCTTT GATGCCCTTG GGGGCTTGAT TGTGGTATAA  120

GGTAGGTTCA GMHAACTCAT TTAAGKTCGC TTTTCCAGGA GTCACTGGGG ACCAGGAATG  180

AGTCCTGGTG CATGGTAATC CCATGCAGAG TTCCCAGCCA TCCCTCTGTC AGCCCAGCAC  240

CTGTGTCTTC CCACATCCAC TGTCAATGCC TTCCCTCTGA GATTTGCTCA GAGTGCGCCT  300

GTCTTCGCA ATG TCC CCA TCT CTT GGT GAC AGA TGT TCC TCT TGG CTG CAT  351
          Met Ser Pro Ser Leu Gly Asp Arg Cys Ser Ser Trp Leu His
                      -40                 -35                 -30

CTA GTC AGC CAT CTT GAA TCA ATC TCT GGA CCT CTA CTT AAT ATC CCT      399
Leu Val Ser His Leu Glu Ser Ile Ser Gly Pro Leu Leu Asn Ile Pro
                -25                 -20                 -15

GAG AAT CTC CTT CTC TGT TGT CAC CGG TGC ACC AAC TGC CAA AGA CAC      447
Glu Asn Leu Leu Leu Cys Cys His Arg Cys Thr Asn Cys Gln Arg His
            -10                  -5                   1

CAT TTT TGC AGT GTT TGG                                              465
His Phe Cys Ser Val Trp
    5
```

(2) INFORMATION FOR SEQ ID NO: 105:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 141 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 1..72
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.4

seq YFLLPCLINLAIG/VK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 105:

```
ATG TCA GGA GCT GAG CCA ACA ACC TTC ATC AGA TAT TTT CTT CTT CCA      48
Met Ser Gly Ala Glu Pro Thr Thr Phe Ile Arg Tyr Phe Leu Leu Pro
                -20                 -15                 -10

TGT CTC ATA AAT TTG GCC ATT GGA GTT AAA TGG AAA ACA GCT TGG AAG      96
Cys Leu Ile Asn Leu Ala Ile Gly Val Lys Trp Lys Thr Ala Trp Lys
            -5                   1               5

AGG GGG GAA AGG CAG CTA AAC AAC ACT GTC TTT TTT TTT TTT TTT         141
Arg Gly Glu Arg Gln Leu Asn Asn Thr Val Phe Phe Phe Phe Phe
     10                  15                  20
```

(2) INFORMATION FOR SEQ ID NO: 106:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 186 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 91..156
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.4
seq QLLFSFLLSTIPT/SY

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 106:

```
ATGTTACTAG GCAATAGGAG TTTTTGAACT GCATTATAAT CTTATGAGAC CACCATCATA    60

TATACAGTTT GTTGTTGACC AAAACGTTAT ATG GTG TAT GAC TAT TTT ATT TCC    114
                                 Met Val Tyr Asp Tyr Phe Ile Ser
                                             -20             -15

CAA CAA CTG CTG TTC TCT TTT TTA CTC TCT ACT ATC CCC ACA TCT TAC     162
Gln Gln Leu Leu Phe Ser Phe Leu Leu Ser Thr Ile Pro Thr Ser Tyr
                -10                 -5                   1

CAC CTT TCC CTT ACT TGC CAG CGG                                     186
His Leu Ser Leu Thr Cys Gln Arg
         5                  10
```

(2) INFORMATION FOR SEQ ID NO: 107:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 332 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 261..302
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.3
seq LFLCSCSLSLNQL/LT

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 107:

```
AAATCAGATT CACCTAAAAC GGAGATCATT TATAGAAAGA AAAAAACACA AAGGTCTTCT    60

GTTGAACTGT TCAGTAAGTG TGTACTGATA CAATAATATG GCAGGATATT GGGTTTCCAA   120

CTCAGAGTTG ATAGTTCAGA ATTAAATTTT GGCCTATGTC TTGCTTTTTA GCTTTTTAAG   180

CCATTCTGAC GCTACCATAG ACATGCCCAT CTGGTAGATG AGAAATTCAG GTGTAGAAAG   240

ACTTGCAGAA ATCTTATGAG ATG CTC TTT CTT TGT TCC TGT TCT CTT TCT CTG   293
                      Met Leu Phe Leu Cys Ser Cys Ser Leu Ser Leu
                                      -10                 -5

AAC CAG CTC CTA ACT TAC ATC TTT GTA GTC CCA CCC TGG                 332
Asn Gln Leu Leu Thr Tyr Ile Phe Val Val Pro Pro Trp
             1                5                  10
```

(2) INFORMATION FOR SEQ ID NO: 108:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 187 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 122..166
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.2

seq FLMVLLFRSNKWT/GK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 108:

```
AAAGAAAATG AACCCCTGTA AAGCAATTTC AGCTTATGAA CCCCTGTAAA GCAATTTCAG   60

CTTTGCTTCA TGGTATATTT GCTTTCTATC CAGAGGTTAA TTTGTTAGTT TTCTTAAAAC  120

A ATG TTT TTC TTA ATG GTC TTA TTG TTT AGA AGT AAC AAG TGG ACT GGA  169
  Met Phe Phe Leu Met Val Leu Leu Phe Arg Ser Asn Lys Trp Thr Gly
  -15                 -10                      -5                1

AAA GTA TAC GGG GCC CTG                                            187
Lys Val Tyr Gly Ala Leu
              5
```

(2) INFORMATION FOR SEQ ID NO: 109:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 200 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 78..164
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.2
seq FLSHVTTSLASSS/AP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 109:

```
ACAACTACCA CTCTCTGTTC TGTTCACTCC GTTCCAGCCA CACCCACCTT CTTGCTGTTC   60

TTTGAACATG GCCTGGC ATG CTC CCT CTT CAG GGC CTT TGC ACT TGT TAT    110
                   Met Leu Pro Leu Gln Gly Leu Cys Thr Cys Tyr
                                   -25                 -20

TTC CTC CAC CTA GAA TTT CTT TCC CAT GTA ACT ACC TCA CTT GCT TCA   158
Phe Leu His Leu Glu Phe Leu Ser His Val Thr Thr Ser Leu Ala Ser
            -15                 -10                 -5

TCA TCA GCT CCC TCA CCT AAA CCT TCA GTA ACC CTT TCT TCG           200
Ser Ser Ala Pro Ser Pro Lys Pro Ser Val Thr Leu Ser Ser
         1               5                  10
```

(2) INFORMATION FOR SEQ ID NO: 110:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 276 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 205..267
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.2
seq HFFLLLNTILLFG/CA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 110:

```
ACAGAGAATT GTTCCACTAC ACTAAAAATC CCATGTGCTC TGCTTATTCA TCCCTCCTTC   60

TCTCCCTCTA GCCCCTGACA ACCACTGATG TCTTTACTGT CTCCCTAGTT TTGCTTTGCC  120

CAGAATGTTA TATAGATGGA ATAATATAGT ATATATTTTC ACATTGGCTT CATTCACTTA  180

GATACATGTC TTTAAGGTTC CTTC ATG TAT TTT TAT GGC TTG ACA TTT CAT     231
                           Met Tyr Phe Tyr Gly Leu Thr Phe His
                               -20                 -15

TTC TTC TTA TTG CTG AAT ACT ATT TTA TTG TTT GGG TGT GCC CGG        276
Phe Phe Leu Leu Leu Asn Thr Ile Leu Leu Phe Gly Cys Ala Arg
    -10                      -5                   1
```

(2) INFORMATION FOR SEQ ID NO: 111:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 271 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 152..238
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.1
seq LWASQGSLQDAQS/ER

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 111:

```
ACTCTATTAT CTCGGCTTCT CGGGAGGAGC CTCATCTAGT CAGTCACGCA GAAGTTTCTC   60

TTTCGCTCTT CGCGCTACAC ACCCAGATTG GCTTCCAGCG CGCAGGTAAA ACCTGGCTGT  120

GCCCTGTTGA AATCAATTCT GTTGCAGTCA T ATG CGG TGG AAT CTG TTC TTC     172
                                   Met Arg Trp Asn Leu Phe Phe
                                                   -25

TTT TGC ATC CTA CGT AAC CAG ACC AAG CTG TGG GCT TCT CAA GGT AGC    220
Phe Cys Ile Leu Arg Asn Gln Thr Lys Leu Trp Ala Ser Gln Gly Ser
        -20                 -15                 -10

CTC CAG GAT GCA CAG AGT GAG AGA GGA TGC TTT TCC CTA AAC CAG GAT    268
Leu Gln Asp Ala Gln Ser Glu Arg Gly Cys Phe Ser Leu Asn Gln Asp
     -5                  1               5                  10

GGG                                                                271
Gly
```

(2) INFORMATION FOR SEQ ID NO: 112:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 245 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 150..191
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.1
seq FIAALFTMAKTWN/QP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 112:

```
AGTGTAAATT AGTTCAACCA TTGTGGAAGA TAGTGTGGCG ATTCCTCAAG GATCTAGAAC   60

CAGAAATACC ATTTGACCCA GCAATCCCAT CATTGGGTAT ATACCCAAAG GATTATAAAT  120

CATTCTAATA TAAAAACACG TGTACGCAT ATG TTT ATT GCA GCA CTA TTC ACA    173
                                Met Phe Ile Ala Ala Leu Phe Thr
                                                -10

ATG GCA AAG ACT TGG AAC CAA CCC GGG TGC TCA TCA ATG ATG GGC TGG    221
Met Ala Lys Thr Trp Asn Gln Pro Gly Cys Ser Ser Met Met Gly Trp
     -5                  1               5                  10

ATA AAG AAA ATG AGG CAC ATG ACG                                    245
Ile Lys Lys Met Arg His Met Thr
```

15

(2) INFORMATION FOR SEQ ID NO: 113:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 207 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 25..156
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.1
seq LWVXLPXXXVIAS/VV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 113:

```
ATTACATYAA ATKTATCAAA ATGY ATG CCY GGG YCA AAG CAT TTT CTC AGA      51
                          Met Pro Gly Xaa Lys His Phe Leu Arg
                                          -40

GTG TTC AGA AMA TCC GCG MKG AGG AGC GTC GGA TAT KGG MAM AAG CCG      99
Val Phe Arg Xaa Ser Ala Xaa Arg Ser Val Gly Tyr Xaa Xaa Lys Pro
-35                 -30                 -25                 -20

GGT ACT TCC AGA GCA TCA CTG TGG GTR TSG CTC CCA TTS RTG GYS GTG     147
Gly Thr Ser Arg Ala Ser Leu Trp Val Xaa Leu Pro Xaa Xaa Xaa Val
                -15                 -10                  -5

ATY GCC AGC GTG GTG ACC TTC TCT GKK CAT ATG ACC CTG GGC TTC GAT     195
Ile Ala Ser Val Val Thr Phe Ser Xaa His Met Thr Leu Gly Phe Asp
              1               5                  10

CTG ACA GCA GCG                                                     207
Leu Thr Ala Ala
     15
```

(2) INFORMATION FOR SEQ ID NO: 114:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 269 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Surrenals

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 30..176
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.1
seq VALGPLFVTGHFA/GE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 114:

```
AAGGCTCCCC TTTCGGCCTC TCGTTCTTT ATG CGA CTT GAA TCT CCC GAC GAA        53
                                Met Arg Leu Glu Ser Pro Asp Glu
                                                -45

AAT TTT GCG GTA GTT CAG GAA CAT GCC ATT CAT CAC ATC GAT GGC CCG       101
Asn Phe Ala Val Val Gln Glu His Ala Ile His His Ile Asp Gly Pro
    -40                 -35                 -30

CTC CGC AGA TTC CTG CTT TTG GAA GTG CAC GAA CCC GTA GCC CTT GGG       149
Leu Arg Arg Phe Leu Leu Leu Glu Val His Glu Pro Val Ala Leu Gly
-25                 -20                 -15                 -10

CCC CTT TTC GTC ACA GGC CAC TTT GCA GGA GAG GAT GTT GCC GAA CGC       197
Pro Leu Phe Val Thr Gly His Phe Ala Gly Glu Asp Val Ala Glu Arg
                 -5                   1               5

CGA GAA GAT GTT GTA CAG CGC CTT GTT GTC GAT GGT CTT GCC CAG GTT       245
Arg Glu Asp Val Val Gln Arg Leu Val Val Asp Gly Leu Ala Gln Val
         10                  15                  20

CTT GAT GAA GAC GTT GCC CAC CCG                                       269
Leu Asp Glu Asp Val Ala His Pro
     25                  30
```

(2) INFORMATION FOR SEQ ID NO: 115:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 249 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 175..234
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4
seq EVLLPTVLRGSYC/FS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 115:

```
ACATTAATCA ACTGTGAAAT ACAGAGCAGG TCACTTCACC TCTCAGTGTG TCCTCATTTT   60

AAAAATCAGA CCGTAACAGT AGCTATCTCA WTAGGGTTGT TAGGAGGTGT ACTGTATTAG  120

GATGTTAGGC CTTATACAWN AGAAGAAAAC GGAACAGTGA CGTAAACAAA TTTG ATG    177
                                                            Met
                                                            -20

GCA GGG AGT CCA GAT AGG GAG GTT CTG CTC CCG ACA GTC CTC AGA GGC    225
Ala Gly Ser Pro Asp Arg Glu Val Leu Leu Pro Thr Val Leu Arg Gly
                -15                 -10                 -5

TCA TAT TGT TTC TCC CAC CAT GGG                                    249
Ser Tyr Cys Phe Ser His His Gly
            1               5
```

(2) INFORMATION FOR SEQ ID NO: 116:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 198 base pairs
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: DOUBLE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
- (A) ORGANISM: Homo Sapiens
- (D) DEVELOPMENTAL STAGE: Fetal
- (F) TISSUE TYPE: brain

(ix) FEATURE:
- (A) NAME/KEY: sig_peptide
- (B) LOCATION: 10..165
- (C) IDENTIFICATION METHOD: Von Heijne matrix
- (D) OTHER INFORMATION: score 4
  seq RHLFLFEISLVFS/KE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 116:

```
AATGATGCC ATG CAT GTC AGC ATG CTG GAA GGG TTC GAC GAG AAC CTG GAT    51
          Met His Val Ser Met Leu Glu Gly Phe Asp Glu Asn Leu Asp
                  -50                 -45                 -40

GTG CAG GGG GAG TTG ATT CTC CAG GAT GCC TTT CAA GTG TGG GAC CCG     99
Val Gln Gly Glu Leu Ile Leu Gln Asp Ala Phe Gln Val Trp Asp Pro
            -35                 -30                 -25

AAG TCG CTG ATC CGG AAG GGG CGG GAG CGG CAC TTG TTC CTC TTT GAG    147
Lys Ser Leu Ile Arg Lys Gly Arg Glu Arg His Leu Phe Leu Phe Glu
        -20                 -15                 -10

ATC TCC TTG GTT TTT AGC AAG GAG ATC AAA GAT TCT TCA GAA CAC AAC    195
Ile Ser Leu Val Phe Ser Lys Glu Ile Lys Asp Ser Ser Glu His Asn
    -5                      1               5                  10
```

```
GGG                                                                   198
Gly
```

(2) INFORMATION FOR SEQ ID NO: 117:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 306 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 112..228
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.9
seq VLAIGLLHIVLLS/IP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 117:

```
AGCAGCCGCG GGTTGTTACA GCTGCTGGAG CAGCAGCGGC CCCCGCTCCC GGGAACCGTT    60

CCCGGGCCGT TGATCTTCGG CCCCACACGA ACAGCAGAGA GGGGCANMAG G ATG AAT   117
                                                         Met Asn

GTG GGC ACA GCG CAC AGC GAG GTG AAC CNC AAC ACG CGG GTG ATG AAG    165
Val Gly Thr Ala His Ser Glu Val Asn Xaa Asn Thr Arg Val Met Lys
        -35                 -30                 -25

GNC CGT GGC ATC TGG CTS TCC TAC GTG CTG GCC ATC GGT CTC CTC CAC    213
Xaa Arg Gly Ile Trp Leu Ser Tyr Val Leu Ala Ile Gly Leu Leu His
    -20                 -15                 -10

ATC GTG CTG CTG AGC ATC CCG TTT GTG AGT GTC CST GTC GTC TGG ACC    261
Ile Val Leu Leu Ser Ile Pro Phe Val Ser Val Xaa Val Val Trp Thr
-5                      1               5                  10

CTC ACC AAC CTC ATT CAC AAC ATG GGC ATG TAT ATC TTC CTG CAC        306
Leu Thr Asn Leu Ile His Asn Met Gly Met Tyr Ile Phe Leu His
            15                  20                  25
```

(2) INFORMATION FOR SEQ ID NO: 118:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 433 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 371..418
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.8
seq FVXAIXXYIPTNS/VQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 118:

```
ATTTCTTCTA TGTCTTGTTT ATTTCACTTA GCATAATGTC CTCTAGGTTC ACTCATGCTG    60

TCACAACTTG GCAGGATTTT CTTCTTTTTT AAGGCTTGAA AATATTCCAT TTTGTGTGTA   120

TATGTGTATG TATGTATACA TACACACACA CATACACAAG CACACATACA CTATAATTTC   180

TTAATCCATT CAACTATTGA TCGATACTTA AATTGATTCC AAATCTTGGC TATTGTGAAT   240

AATGCTGCAA TGAACATGGG AGTACAGATA TCTCTTCAAC ATACTGAGTT CAAATCTTTT   300

GGGTAAATAC CCAGATGTGG GATTGCTGGA TCATATGATA ATTCTATTTT TAATTTTTTG   360

AGTGACVTCC ATG CTA TCA TTC GTA NTG GCT ATA KDA NTT TAC ATT CCT      409
           Met Leu Ser Phe Val Xaa Ala Ile Xaa Xaa Tyr Ile Pro
               -15                 -10                 -5

ACT AAT AGT GTA CAA GAA TTC CTT                                     433
Thr Asn Ser Val Gln Glu Phe Leu
            1               5
```

(2) INFORMATION FOR SEQ ID NO: 119:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 403 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 284..379
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.8
seq TFINITLWLGSLC/QR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 119:

```
ACAGCTGGGG CTTTGTCTTC TTTATTGCTA GGAGAATGTA GCAATAGAAG TTCTCATCGC   60

CCTGTATTGC ACTTTTGGTT TTAAGGACTG GACCCAGAGT TCCTGAAAGC CAAACTCCAT  120

AAGCTGCTCA GTAAGTTCCA AGCACATAGC CGGCTKHGGG ATGCGATTCG GTCGAGGTCT  180

GTTGAATGAA GGTAGACGCA GCAGGCAGTT TGTCCTTACC AGTGACCTGG AAGACGGTGG  240

CACTTCCTGA GTGAGCTCAC TTACCTTCCC TGAATGGTGA GGC ATG GAT GAA TAT    295
                                                Met Asp Glu Tyr
                                                        -30

TCC TGG TGG TGC CAC GTG TTA GAG GTG GTA AAG GGT CAA ATG TTT ACT    343
Ser Trp Trp Cys His Val Leu Glu Val Val Lys Gly Gln Met Phe Thr
            -25                 -20                 -15

TTT ATT AAT ATT ACA TTA TGG CTT GGT TCT CTG TGT CAG CGA TTT TTC    391
Phe Ile Asn Ile Thr Leu Trp Leu Gly Ser Leu Cys Gln Arg Phe Phe
        -10                 -5                   1

TAT GCC TCG GGT                                                    403
Tyr Ala Ser Gly
  5
```

(2) INFORMATION FOR SEQ ID NO: 120:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 181 base pairs
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: DOUBLE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
- (A) ORGANISM: Homo Sapiens
- (F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
- (A) NAME/KEY: sig_peptide
- (B) LOCATION: 95..163
- (C) IDENTIFICATION METHOD: Von Heijne matrix
- (D) OTHER INFORMATION: score 3.8
  seq FIFLIQIWKTCLS/FS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 120:

```
ACATGATGGA TGAATGATGT TATCTTTCAC ATGTGGCTTT CTTGCCTTGT CCTAAGTGCC   60

TGCTGTAGTC GTTGACATTT TCCAGCAAAC AGGA ATG AGG AGA AAA GGT CAA GGA  115
                                      Met Arg Arg Lys Gly Gln Gly
                                                  -20

CAT CTA GCC TTT ATC TTC CTG ATT CAG ATT TGG AAA ACA TGC CTT TCG    163
His Leu Ala Phe Ile Phe Leu Ile Gln Ile Trp Lys Thr Cys Leu Ser
    -15                 -10                 -5
```

```
TTT TCT CCC ACC TCT GGG                                              181
Phe Ser Pro Thr Ser Gly
  1               5
```

(2) INFORMATION FOR SEQ ID NO: 121:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 129 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 37..111
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.8
seq NILFLAVSSFSMP/LP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 121:

```
AATATTGTAA GGTACCTTAT TCCTTAGTGG TATGGA ATG TTT TTA ATC AGT GGA      54
                                        Met Phe Leu Ile Ser Gly
                                        -25                 -20

CAT GTG CAT TTA ATT TAT AAC ATC CTG TTC CTG GCA GTA TCG TCT TTT     102
His Val His Leu Ile Tyr Asn Ile Leu Phe Leu Ala Val Ser Ser Phe
                -15                 -10                 -5

TCC ATG CCC CTG CCC TGC CTC TAC AGG                                 129
Ser Met Pro Leu Pro Cys Leu Tyr Arg
            1               5
```

(2) INFORMATION FOR SEQ ID NO: 122:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 332 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 99..290
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.8
seq LFIVVCVICVTLN/FP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 122:

```
ATTGGATTAG TAGAATTGCT TTTGTCATTC CATTGTTTTC ATATATTTGT TTGGGACATT    60

TTACTTTTTT CTGTTAACGC TTACCCTAGR AATTAGAA ATG ACA CCA CGT ATT CTT   116
                                          Met Thr Pro Arg Ile Leu
                                                          -60

AGC GAA GTC CAG TTT TCA GCA TTT TGT CCT TAT TGG ACA ATA GCA AGG    164
Ser Glu Val Gln Phe Ser Ala Phe Cys Pro Tyr Trp Thr Ile Ala Arg
            -55                 -50                 -45

ATA TTA GAA CGT GTT GGT TCC GCG TGC TTC CGT CTT GAG TTA TGT GCT    212
Ile Leu Glu Arg Val Gly Ser Ala Cys Phe Arg Leu Glu Leu Cys Ala
        -40                 -35                 -30

GCT ATT GTC GGA TAT TTT GTC TTA GAT GTA CGT ACT TTC CTG TTC ATT    260
Ala Ile Val Gly Tyr Phe Val Leu Asp Val Arg Thr Phe Leu Phe Ile
    -25                 -20                 -15

GTG GTA TGT GTA ATT TGC GTT ACT TTG AAT TTT CCA CGT THN KAC TTT    308
Val Val Cys Val Ile Cys Val Thr Leu Asn Phe Pro Arg Xaa Xaa Phe
-10                     -5                   1               5

CTT TGT CTC TCA TCA CTT ACC GCG                                    332
Leu Cys Leu Ser Ser Leu Thr Ala
            10
```

(2) INFORMATION FOR SEQ ID NO: 123:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 225 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 25..66
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq CSLLSGWGQLLRC/VQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 123:

```
AAGTGGTCCC AGGACCACTC ATGG ATG TGT AGT TTG CTG AGT GGC TGG GGA       51
                          Met Cys Ser Leu Leu Ser Gly Trp Gly
                                          -10

CAG CTT CTT AGA TGT GTA CAG ACC CCA GCA GAG CCC AGA GAT GTA AAC      99
Gln Leu Leu Arg Cys Val Gln Thr Pro Ala Glu Pro Arg Asp Val Asn
 -5                  1               5                  10

AAG AAG CBA GAA AAA AAG GAA AAG TAC ATG CCC CTG GTA GAT TCC CTC     147
Lys Lys Xaa Glu Lys Lys Glu Lys Tyr Met Pro Leu Val Asp Ser Leu
            15                  20                  25

TGT GGA GGG TTA GGA ACT AGG AAT AGT GAC TGC CTC AGA GGG GGA GCT     195
Cys Gly Gly Leu Gly Thr Arg Asn Ser Asp Cys Leu Arg Gly Gly Ala
        30                  35                  40

GGC AGG GGC AGA GAT GGG AGA AGG ATA AGG                             225
Gly Arg Gly Arg Asp Gly Arg Arg Ile Arg
    45                  50
```

(2) INFORMATION FOR SEQ ID NO: 124:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 281 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 144..254
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq LIPFNFSASGLCA/CS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 124:

```
ATATTGTCAC TAAATTTAGT GGCCATGTCT CCATCCTTCT AACCTGTCTG GCCTTCTTAG   60

TAATACATTG CATGGGCTAA CTTGTTTTTT TTGGAAMAWG CMSATVSATT KGAACKKCTT  120

CAAATCAGTA CCCCCTGGCC TCV ATG CTG TTT TCT TTC TGC TTC CCA GTT CAT  173
                          Met Leu Phe Ser Phe Cys Phe Pro Val His
                                  -35                 -30

TTC TGG AAC CCA AGT TCA CTC TTT CCC CCA AGT TCA GTT TCT TTG ATC     221
Phe Trp Asn Pro Ser Ser Leu Phe Pro Pro Ser Ser Val Ser Leu Ile
        -25                 -20                 -15

CCC TTT AAT TTC AGT GCA TCT GGG CTG TGT GCT TGT TCT AGR ACC TTC     269
Pro Phe Asn Phe Ser Ala Ser Gly Leu Cys Ala Cys Ser Arg Thr Phe
```

```
    -10                    -5                    1              5

ACA CAC ATG GGT                                                        281
Thr His Met Gly
```

(2) INFORMATION FOR SEQ ID NO: 125:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 364 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 284..328
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.6
seq WILRILFVIGSXL/EK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 125:

```
ATGACTGCCA TTTGAAGTGG CTACTTCAAT GGTTGGTTGA TAATAACTTT CAACATTCTG    60

TGAATGTAAG CTGTGCACAC CCTGAATGGC TAGCAGGGCA AAGCATCCTG AATGTGGATC   120

TGANAGATTT WRTCTGTGAT GATTTTCTCA AGCCACAGAT AAGGACACAT CCWGAAACCA   180

TAATTGCTCT AAGAGGCATG ANTGTGACTC TGACGTGCAC TGCAGTGAAG CAGMAGTGAT   240

TCACCCATGT CCACTGTGTG GCGCAAAGAC AGTGAAATCC TGT ATG ACG TGG ATA    295
                                                Met Thr Trp Ile
                                                -15

CTG AGA ATT TTG TTC GTT ATT GGC AGC ARD CTG GAG AAG CTC TGG AAT    343
Leu Arg Ile Leu Phe Val Ile Gly Ser Xaa Leu Glu Lys Leu Trp Asn
    -10                 -5                  1                  5

ATA CTA GTA TCT TAC ATC TTT                                        364
Ile Leu Val Ser Tyr Ile Phe
                10
```

(2) INFORMATION FOR SEQ ID NO: 126:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 290 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 123..266
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.6
seq ILCIFLGLLIIRC/FK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 126:

```
AGACAAACGT GCCAACACTT AAGTCTACTG GCTGGACTTC ATCTCCATGG CAACAAGCAT   60

GGAAGGCAAA GAGTTGATTC CAGAAGGAAC TGTGAAGAGC CACAACAATG TGCCAGTGAA  120

TA ATG AGT AGT ACC TAC TGT GGC AAC TCT TCA GCT AAG ATG AGT GTC      167
   Met Ser Ser Thr Tyr Cys Gly Asn Ser Ser Ala Lys Met Ser Val
               -45                 -40                 -35

AAC GAA GTA TCA GCT TTC TCA TTG AGT CTG GAG CAA AAA ACT GGC TTT     215
Asn Glu Val Ser Ala Phe Ser Leu Ser Leu Glu Gln Lys Thr Gly Phe
            -30                 -25                 -20

GCT TTT GTT GGG ATT TTG TGT ATC TTC TTG GGA CTT CTT ATT ATC CGA     263
Ala Phe Val Gly Ile Leu Cys Ile Phe Leu Gly Leu Leu Ile Ile Arg
        -15                 -10                  -5

TGC TTC AAA ATC CTG CTA GNS CAA TCG                                 290
Cys Phe Lys Ile Leu Leu Xaa Gln Ser
      1               5
```

(2) INFORMATION FOR SEQ ID NO: 127:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 208 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 143..196
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.6
seq LTMLSMIVGATCY/AM

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 127:

```
AGCTGCAGGA CTTCCCGCGC AACTGCTGGG TGTCCATCAA TGGCATGGTG AACCACTCGT   60

GGAGTTAACT GTACTCCTTC GCACTCTTCA AGGCCATGAG CCACATGCTG TGCATCGGGT  120

ACGGCCGGCA GSGCCCGAGA GC ATG ACG GAC ATC TGG CTG ACC ATG CTC AGC  172
                         Met Thr Asp Ile Trp Leu Thr Met Leu Ser
                                     -15                 -10

ATG ATT GTG GGT GCC ACC TGC TAC GCC ATG ATC GGG                     208
Met Ile Val Gly Ala Thr Cys Tyr Ala Met Ile Gly
            -5                      1
```

(2) INFORMATION FOR SEQ ID NO: 128:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 214 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 59..109
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.6
seq WIYAFISLGYILG/SG

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 128:

```
AAGTTTGGGT TGTTTCTGCT TTTCGTTAAA AATAGTAGTA CTTCTCTGAA CACTGTGG      58

ATG AGH NTT TGC TGG ATA TAT GCT TTC ATT TCT CTT GGG TAT ATA CTT     106
Met Xaa Xaa Cys Trp Ile Tyr Ala Phe Ile Ser Leu Gly Tyr Ile Leu
    -15                     -10                 -5

GGG AGT GGA ATT GTT GGG TTA TTT GGT AAT TTT ATG TTT AAA CTT TTG     154
Gly Ser Gly Ile Val Gly Leu Phe Gly Asn Phe Met Phe Lys Leu Leu
      1               5                  10                  15

AGG AAC TGC CAG ACC GTT TTC CAG GAT GGC TAT GCT ATA TTA CCC TTC     202
Arg Asn Cys Gln Thr Val Phe Gln Asp Gly Tyr Ala Ile Leu Pro Phe
                 20                  25                  30

CCA CCA ACG GGG                                                     214
Pro Pro Thr Gly
         35
```

(2) INFORMATION FOR SEQ ID NO: 129:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 156 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 19..111
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq QLALSWVPPXCRV/SP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 129:

```
ACTTTTAGGC CCATTGGG ATG TTC ATT AGA ACT CTG AAA ACT ACA GTT CTC      51
                    Met Phe Ile Arg Thr Leu Lys Thr Thr Val Leu
                        -30                 -25

CCC TTT ATG AGG ACT GCA CCA CAG CTC GCC CTC TCC TGG GTT CCG CCT      99
Pro Phe Met Arg Thr Ala Pro Gln Leu Ala Leu Ser Trp Val Pro Pro
-20                 -15                 -10                  -5

RRT TGC AGA GTG AGC CCA TGG GAC AGC CCT CTG AAA TTA TAC TGC TTA     147
Xaa Cys Arg Val Ser Pro Trp Asp Ser Pro Leu Lys Leu Tyr Cys Leu
                  1                   5                  10

CAA CCG CAG                                                         156
Gln Pro Gln
         15
```

(2) INFORMATION FOR SEQ ID NO: 130:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 297 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 211..282
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5

seq SVLIFCLLPYIYH/FF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 130:

```
ATAAAAATAA TTTATATTTG ATGTTATAAT ACATTATTTA ATTTTTATAG TAAATTCACT    60

AATCTGTTTT GTTTAGACCA GKRTGAGAGA AAAGSSAGAC TKCGAGTTAA TACTTTGTAA   120

GGTTATCTGC ACTCTCATCT GTGGTTGGCA ATATTTGATG CAGTTTATAT TAGATTTATG   180

TTGTTGTTAT TTAGATAGTT TGGAGCTGAG ATG AGG ACA GGA GCT GAG ATG AGG     234
                                 Met Arg Thr Gly Ala Glu Met Arg
                                                 -20

ACA AAC TCT TCA GTT TTA ATT TTT TGT TTG CTC CCA TAT ATT TAT CAT     282
Thr Asn Ser Ser Val Leu Ile Phe Cys Leu Leu Pro Tyr Ile Tyr His
    -15                 -10                 -5

TTT TTT CCA GCT TCG                                                 297
Phe Phe Pro Ala Ser
 1               5
```

(2) INFORMATION FOR SEQ ID NO: 131:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 310 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 146..214
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq EGLELGFSHRTFA/FP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 131:

```
ATAGGGCAAT TANNGGCTCA TCTCATTTGT TTCCWNWNCT CTCAGGGGCC ACAGCCCTGC    60

ATTACCTGTT TGCCAATGTG TGAAAGCAAT TGTTTCATGT ATTTTGCCCA GTTTTCTAGT   120

TTCTTATGGC AGGAAGTTAA GTCCA ATG ATT GTT ATT CCA TCA TGG CTG GAA     172
                            Met Ile Val Ile Pro Ser Trp Leu Glu
                                        -20                 -15

AAC GAG GGT CTG GAA TTG GGA TTT TCA CAT AGG ACC TTT GCT TTT CCT     220
Asn Glu Gly Leu Glu Leu Gly Phe Ser His Arg Thr Phe Ala Phe Pro
                -10                 -5                   1
```

```
GTG ACA CAT GCT TCC TCT CAG TAC ATA TGG ATG AAC TNK CTG ACT AGG      268
Val Thr His Ala Ser Ser Gln Tyr Ile Trp Met Asn Xaa Leu Thr Arg
            5                  10                  15

ACT ACA GTA GCA ATA TCA GTT TAT TTT TGG ACC CAC ACA GGG              310
Thr Thr Val Ala Ile Ser Val Tyr Phe Trp Thr His Thr Gly
     20                  25                  30
```

(2) INFORMATION FOR SEQ ID NO: 132:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 409 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 266..394
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq VISVFLSFLPSYP/GF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 132:

```
ACAAATCTAA ATCTAAAGAT AATTTTCAAG TTTGGCCAGG GAATCCTTAT GTTATATTGA   60

TTCCATAAAA ACAGGGGTAT GTATTACCTA GCAGCTGAAG AAGAGATGGC CTACATTAGG  120

CACCTGAGTG TCAGAAGGGG TTGTCAAGAA AGTGTACAGT TAATACATAA GTCTGTCCTG  180

ATTGAATTAA TTAATCAAGA ACTACAGGTG TTGAAAAGAG AAATGGGTAC CTGGAGAAGG  240

GCATCATCTA TATTGGCATC TTGGA ATG TTA AAA AAG GAA ATA GCT CAC CAC     292
                            Met Leu Lys Lys Glu Ile Ala His His
                                        -40                 -35

AGC CCT AGC CTG GTG AGC TGC CCT GTC TGC ACC ACA AAA TAT AGA ACT      340
Ser Pro Ser Leu Val Ser Cys Pro Val Cys Thr Thr Lys Tyr Arg Thr
                -30                 -25                 -20

CTG AGA CTC CTG AGG GTT ATC TCA GTT TTT CTG TCT TTT CTT CCT TCT      388
Leu Arg Leu Leu Arg Val Ile Ser Val Phe Leu Ser Phe Leu Pro Ser
            -15                 -10                  -5

TAC CCA GGG TTC AGC ATG CAA                                          409
Tyr Pro Gly Phe Ser Met Gln
         1                  5
```

(2) INFORMATION FOR SEQ ID NO: 133:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 343 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 112..344
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 36..268
id AA256780
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 248..337
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq CAYSLPGVALTLG/RQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 133:

```
TGATCGGACC ATTTCACTGC AGCAAGCAAC ACAGTATTCT GAGCAGAAGA TCGGGACTTG   60

AGGCCATGTT GCGGAGGGCC AGTGACATTA TCTGGACTCT GGAGTGTGRR GRAATATBGA  120

STCCACKCTT CACTATATTC ACAGCGATTC AGACTTGAGC AACAATAGCA GTTTTAGCCC  180

TGATGAGGAA AGGAGAACTR RAGTACAAGA TGTTGTACCT CAGGCGTTGT TAGATCAGKA  240

TTTATCT ATG ACT GRS CCT TCT CGT GCA CAG ACG GTT GAC ASK GGA ATT    289
        Met Thr Xaa Pro Ser Arg Ala Gln Thr Val Asp Xaa Gly Ile
        -30                 -25                 -20

GCT AAG CAC TGT GCA TAT AGC CTC CCT GGT GTG GCC TTG ACA CTC GGA    337
Ala Lys His Cys Ala Tyr Ser Leu Pro Gly Val Ala Leu Thr Leu Gly
    -15                 -10                  -5

AGA CAG                                                            343
Arg Gln
  1
```

(2) INFORMATION FOR SEQ ID NO: 134:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 151 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 52..133
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 55..136
id W81722
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 90..150
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 1..61
id R67182
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 80..114
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 1..35
id R57498
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 20..50
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 297..327
id N21080
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 71..133
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9.8
seq LGLLCALLPQHHG/AP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 134:

```
AGCAGCTCCC AGGATGAACT GGTTGCAGTG GCTGCTGCTG CTGCGGGGGC GCTGAGAGGA    60

CACGAGCTCT ATG CCT TTC CGG CTG CTC ATC CCG CTC GGC CTC CTG TGC      109
           Met Pro Phe Arg Leu Leu Ile Pro Leu Gly Leu Leu Cys
               -20                 -15                 -10

GCG CTG CTG CCT CAG CAC CAT GGT GCG CCA GGT CCC GAY KGG             151
```

```
Ala Leu Leu Pro Gln His His Gly Ala Pro Gly Pro Asp Xaa
            -5                   1               5
```

(2) INFORMATION FOR SEQ ID NO: 135:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 244 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 68..245
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 1..178
id T08712
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 79..245
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 2..168
id R88049
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 80..154
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 1..75
id AA094697
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 170..202
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 89..121
id AA094697
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 112..245
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 91
region 35..168

```
                                    id H30765
                                    est

      (ix) FEATURE:
           (A) NAME/KEY: other
           (B) LOCATION: 79..121
           (C) IDENTIFICATION METHOD: blastn
           (D) OTHER INFORMATION:  identity 97
                                   region 1..43
                                   id H30765
                                   est

      (ix) FEATURE:
           (A) NAME/KEY: other
           (B) LOCATION: 112..245
           (C) IDENTIFICATION METHOD: blastn
           (D) OTHER INFORMATION:  identity 93
                                   region 21..154
                                   id H38484
                                   est

      (ix) FEATURE:
           (A) NAME/KEY: sig_peptide
           (B) LOCATION: 134..187
           (C) IDENTIFICATION METHOD: Von Heijne matrix
           (D) OTHER INFORMATION:  score 9.2
                                   seq VFLCSLLAPMVLA/SA

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 135:

AAGGASGCAG GGGAGGCGGG AAAGCAGCTC AAGCCTCACC CACCGCCCTG CCCCCAGCCC   60

CGCSACTCCC AGGCTCCTCG GGACTCGGCG GGTCCTCCTG GGAGTCTCGS AGGGGACCGG  120

CTGTGCAGAC GCC ATG SAG TTG GTG CTG GTC TTC CTC TGC AGC CTG CTG      169
               Met Xaa Leu Val Leu Val Phe Leu Cys Ser Leu Leu
                           -15                 -10

GCC CCC ATG GTC CTG GCC AGT GCA GCT GAR AAG GAG RAG GAM ATG GAS     217
Ala Pro Met Val Leu Ala Ser Ala Ala Glu Lys Glu Xaa Xaa Met Xaa
    -5                  1               5                  10

CCT TTT CAT TAT GAT TAC CAG ACC CTG                                 244
Pro Phe His Tyr Asp Tyr Gln Thr Leu
                15


(2) INFORMATION FOR SEQ ID NO: 136:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 281 base pairs
          (B) TYPE: NUCLEIC ACID
          (C) STRANDEDNESS: DOUBLE
          (D) TOPOLOGY: LINEAR

     (ii) MOLECULE TYPE: CDNA

     (vi) ORIGINAL SOURCE:
          (A) ORGANISM: Homo Sapiens
```

(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 52..283
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 1..232
id AA111270
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 15..104
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9
seq FFLLLLFRGCLIG/AV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 136:

```
AAGCAACCCT CGAC ATG GCG CTG AGG CGG CCA CCG CGA CTC CGG CTC TGC       50
                Met Ala Leu Arg Arg Pro Pro Arg Leu Arg Leu Cys
                -30                 -25                 -20

GCT CGG CTG CCT GAC TTC TTC CTG CTG CTG CTT TTC AGG GGC TGC CTG       98
Ala Arg Leu Pro Asp Phe Phe Leu Leu Leu Leu Phe Arg Gly Cys Leu
            -15                 -10                 -5

ATA GGG GCT GTA AAT CTC AAA TCC AGC AAT CGA ACC CCA GTG GTA CAG      146
Ile Gly Ala Val Asn Leu Lys Ser Ser Asn Arg Thr Pro Val Val Gln
          1               5                  10

GAA TTT GAA AGT GTG GAA CTG TCT TGC ATC ATT ACG GAT TCG CAG ACA      194
Glu Phe Glu Ser Val Glu Leu Ser Cys Ile Ile Thr Asp Ser Gln Thr
 15                  20                  25                  30

AGT GAC CCC AGG ATC GAG TGG AAG AAA ATT CAA GAT GAA CAA ACC ACA      242
Ser Asp Pro Arg Ile Glu Trp Lys Lys Ile Gln Asp Glu Gln Thr Thr
                 35                  40                  45

TAT GTG TTT TTT GAC AAC AAA ATT CAG GGA GAC TTG GCG                  281
Tyr Val Phe Phe Asp Asn Lys Ile Gln Gly Asp Leu Ala
             50                  55
```

(2) INFORMATION FOR SEQ ID NO: 137:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 461 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

```
     (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: 105..234
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 97
                                  region 1..130
                                  id N77056
                                  est

     (ix) FEATURE:
          (A) NAME/KEY: sig_peptide
          (B) LOCATION: 123..302
          (C) IDENTIFICATION METHOD: Von Heijne matrix
          (D) OTHER INFORMATION:  score 7.2
                                  seq VLLTLLLIAFIFL/II

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 137:

AAAGTAATCT TTATTTCGTC ATTTTTGARA CATAGAAGCC GTAACGGAAG CAAGTGAAAT   60

GCTCAGTCTT AGACGACTGC GTCGTGCTAT GACCGGACTT TTTCTTGAAA GGGGATGACA  120

GC ATG GGA GGC AAT GGC TCC ACA TGT AAA CCC GAC ACT GAA AGA CAA     167
   Met Gly Gly Asn Gly Ser Thr Cys Lys Pro Asp Thr Glu Arg Gln
   -60                 -55                 -50

GGC ACT CTC TCC ACA GCA GCC CCA ACA ACT AGC CCT GCA CCC TGT CTC    215
Gly Thr Leu Ser Thr Ala Ala Pro Thr Thr Ser Pro Ala Pro Cys Leu
-45                 -40                 -35                 -30

TCT AAC CAC CAC AAC AAA AAA CAT TTA ATC CTT GCC TTT TGT GCT GGG    263
Ser Asn His His Asn Lys Lys His Leu Ile Leu Ala Phe Cys Ala Gly
                -25                 -20                 -15

GTT CTA CTG ACA CTG CTG CTG ATA GCC TTT ATC TTC CTC ATC ATA ABB    311
Val Leu Leu Thr Leu Leu Leu Ile Ala Phe Ile Phe Leu Ile Ile Xaa
            -10                  -5                   1

AGC TAC AGA AAA TAT CAC TCC AAG CCC CAG GCC CCA GAT CCT CAC TCA    359
Ser Tyr Arg Lys Tyr His Ser Lys Pro Gln Ala Pro Asp Pro His Ser
     5                  10                  15

GAT CCT CCA GCG VNG CTT TCA TBS ATC CCA GGG GAA ATC ACT TAC CTA    407
Asp Pro Pro Ala Xaa Leu Ser Xaa Ile Pro Gly Glu Ile Thr Tyr Leu
 20                 25                  30                  35

TGC CAG CAC AAC TTT CAA ACT CTY NRN WBW AAA AGA GCA ATC ACT TGG    455
Cys Gln His Asn Phe Gln Thr Leu Xaa Xaa Lys Arg Ala Ile Thr Trp
                40                  45                  50

CTG AGA                                                            461
Leu Arg


 (2) INFORMATION FOR SEQ ID NO: 138:

      (i) SEQUENCE CHARACTERISTICS:
           (A) LENGTH: 380 base pairs
           (B) TYPE: NUCLEIC ACID
```

(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 233..381
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 91
region 157..305
id N78069
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 72..129
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 1..58
id N78069
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 117..180
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 45..108
id N78069
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 173..224
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 100..151
id N78069
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 233..381
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 183..331
id AA022144
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 117..234
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 69..186
id AA022144

est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 89..129
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 42..82
id AA022144
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 242..369
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 170..297
id AA043627
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 89..129
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 21..61
id AA043627
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 117..233
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 194..310
id W68800
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 251..304
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 329..382
id W68800
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 89..129
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 167..207
id W68800
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 117..224
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94

```
                                  region 70..177
                                  id H91637
                                  est

     (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: 251..337
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 94
                                  region 202..288
                                  id H91637
                                  est

     (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: 89..129
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 95
                                  region 43..83
                                  id H91637
                                  est

     (ix) FEATURE:
          (A) NAME/KEY: sig_peptide
          (B) LOCATION: 90..206
          (C) IDENTIFICATION METHOD: Von Heijne matrix
          (D) OTHER INFORMATION:  score 7.2
                                  seq SALAKLLLTCCSA/LR

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 138:

AAACTCTCAA CCCACTTCTC CAGCCAGCGC CCCAGCCCTC CCGCCGCCCG CTCGCAGGTC    60

CCGAGGAGCG CAGACTGTGK CCCTGGCAA ATG GGA ACA GCH GAC AGT GAT GAG     113
                                Met Gly Thr Ala Asp Ser Asp Glu
                                                -35

ATG GCC CCG GAG GCC CCA CAG CAC ACC CAC ATC GAT GTG CAC ATC CAC     161
Met Ala Pro Glu Ala Pro Gln His Thr His Ile Asp Val His Ile His
    -30                 -25                 -20

CAG GAG TCT GCC CTG GCC AAG CTC CTG CTC ACC TGC TGC TCT GCG CTG     209
Gln Glu Ser Ala Leu Ala Lys Leu Leu Leu Thr Cys Cys Ser Ala Leu
-15                 -10                  -5                   1

CGG CCC CGG GCC ACC CAG GCC AGG GGC AGC AGC CGG CTG CTG GTG GCC     257
Arg Pro Arg Ala Thr Gln Ala Arg Gly Ser Ser Arg Leu Leu Val Ala
            5                  10                  15

TCG TGG GTG ATG CAG ATC GTG CTG GGG ATC DTG AGT GCA GTC MTA GGA     305
Ser Trp Val Met Gln Ile Val Leu Gly Ile Xaa Ser Ala Val Xaa Gly
        20                  25                  30

GGA TTT TTC TAC ATC CGC GAC TAS ACC CTC MTC GTC ACC TCG GGA GCW     353
Gly Phe Phe Tyr Ile Arg Asp Xaa Thr Leu Xaa Val Thr Ser Gly Ala
    35                  40                  45

GCC ATC TGG ACA GGG GCT GTG GCT GTG                                 380
Ala Ile Trp Thr Gly Ala Val Ala Val
 50                 55
```

(2) INFORMATION FOR SEQ ID NO: 139:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 237 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 201..234
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..34
id R74380
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 175..216
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.6
seq SLLSFLFARVNLG/SP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 139:

```
ACTATCCTAT CTTCTTTAAG GAAAACTTAC CAAGTATGAT TTGGTCCAAG GATTTCAGGT   60

GGGCTCTCAG TGCTGCTCCC AATATATTAG AGGTCTCCTT CCTTTACTAT TTCCTAACCA  120

GATGTAAAAT TAGCTTTTCC CCCCTTCTAC ATCACCTAAC CCGTTTCTCA TTGG ATG    177
                                                            Met

TCC TTG CTT TCG TTT TTA TTT GCC AGA GTA AAT CTA GGA TCT CCC TTG    225
Ser Leu Leu Ser Phe Leu Phe Ala Arg Val Asn Leu Gly Ser Pro Leu
        -10                 -5                  1

TCT GCC AAC GGG                                                    237
Ser Ala Asn Gly
    5
```

(2) INFORMATION FOR SEQ ID NO: 140:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 375 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 90..376
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 5..291
id R59435
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 206..376
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..171
id R19220
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 199..371
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 1..173
id HSC32F041
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 283..371
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 1..89
id HSC2VA021
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 136..264
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.2
seq WWCCPARLTLTSG/WP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 140:

```
ACTCTTTTTG AAGGTCTCCT TTGCCAGCGC ACACGGCTCC CTGGGCTGGA ATGTCTGTTC    60

ATTCATCCCT GCAGTTGTTT GCGGATGTCC CGGGGGHWAA CGTGAGTYAG TTAATGAAGT   120

CCAAAGCCAA GCCCA ATG GCA AGA AGC CCG CTG CGG AGG AGA GGA AGG CCT    171
                 Met Ala Arg Ser Pro Leu Arg Arg Arg Gly Arg Pro
                             -40                 -35
```

```
ACC TGG AGC CTG AGC ACA CCA AGG CCA GGA TCA CCG ACT TCC AGT TCA      219
Thr Trp Ser Leu Ser Thr Pro Arg Pro Gly Ser Pro Thr Ser Ser Ser
    -30                 -25                 -20

AGG AGC TGG TGG TGC TGC CCC GCG AGA TTG ACC TTA ACG AGT GGC TGG      267
Arg Ser Trp Trp Cys Cys Pro Ala Arg Leu Thr Leu Thr Ser Gly Trp
-15                 -10                  -5                   1

CCA GCA ACA CCA CGA CGT TTT TCC ACC ACA TCA ACC TGC AGT ATA GCA      315
Pro Ala Thr Pro Arg Arg Phe Ser Thr Thr Ser Thr Cys Ser Ile Ala
              5                  10                  15

CCA TCT CGG AGT TCT GCA CAG GAG AGA CGT GTC AGA CGA TGG CCG TGT      363
Pro Ser Arg Ser Ser Ala Gln Glu Arg Arg Val Arg Arg Trp Pro Cys
        20                  25                  30

GCA ACA CAC AGT                                                      375
Ala Thr His Ser
    35
```

(2) INFORMATION FOR SEQ ID NO: 141:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 268 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 40..264
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..225
id T05872
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 64..264
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 2..202
id T34681
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 85..264
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..180
id T05903

est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 157..228
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 25..96
id N56217
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 223..264
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 92..133
id N56217
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 158..214
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq TLLLACHLQLEVG/VV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 141:

```
AATCATTTTG AAGACAGAGG AGCAGCTGTT AAGGCTATGG ATGAAATGAA TGGCAAAGAA    60

ATAGAAGGGG AAGAAATTGA AATAGTCTTA GCCAAGCCAC CAGACAAGAA AAGGAAAGAG   120

CGCCAAGCTG CTAGACAGGC CTCCAGAAGC ACTGCGT ATG AAG ATT ATT ACT ACC     175
                                         Met Lys Ile Ile Thr Thr
                                                             -15

ACC CTC CTC CTC GCA TGC CAC CTC CAA TTA GAG GTC GGG GTC GTG GTG     223
Thr Leu Leu Leu Ala Cys His Leu Gln Leu Glu Val Gly Val Val Val
            -10                  -5                   1

GGG GGA GAG GTG GAT ATG GCT ACC CTC CAG ATT ACT ACG GCA TCG         268
Gly Gly Glu Val Asp Met Ala Thr Leu Gln Ile Thr Thr Ala Ser
      5                  10                  15
```

(2) INFORMATION FOR SEQ ID NO: 142:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 195 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

```
(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 118..191
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 98
                             region 112..185
                             id R60698
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 122..191
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 100
                             region 1..70
                             id H17558
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 113..191
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 92
                             region 229..307
                             id N26943
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 118..191
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 94
                             region 145..218
                             id W24886
                             est

(ix) FEATURE:
     (A) NAME/KEY: sig_peptide
     (B) LOCATION: 76..186
     (C) IDENTIFICATION METHOD: Von Heijne matrix
     (D) OTHER INFORMATION:  score 3.7
                             seq FLGVLALLGYLAV/RP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 142:

AATAAGTTCT CGCGAGACGC AKAAGGCACC GGCTCGAACT GGGGCGGGCC ACTGCCAGGA   60

AAGCAACGCC CCTGA ATG CTT ATG CCG GTG GTT GGT AGA GGA AAT GGA ATT   111
                 Met Leu Met Pro Val Val Gly Arg Gly Asn Gly Ile
                         -35                 -30

CCC CAG ACT GTT TCA GAA TGG CTT CGG TTA TTG CCT TTC CTT GGT GTA   159
Pro Gln Thr Val Ser Glu Trp Leu Arg Leu Leu Pro Phe Leu Gly Val
-25                 -20                 -15                 -10

CTC GCA CTT CTT GGC TAC CTT GCA GTT CGT CCC GGG                   195
Leu Ala Leu Leu Gly Tyr Leu Ala Val Arg Pro Gly
                 -5                   1
```

(2) INFORMATION FOR SEQ ID NO: 143:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 386 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 114..343
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 76..305
id T91418
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 41..118
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 2..79
id T91418
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 150..296
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.6
seq PPFFLCLQCFTRG/FE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 143:

```
AAGGCGCGCG CGACCGGCGG CTCTTTGGCG CGGATTAGGG GGTCTCGGCG AGGGAGTCAT   60

CAAGCTTTGG TGTATGTGTT GGCCGGTTCT GAAGTCTTGA AGAAGCTCTG CTGAGGAAGA  120

CCAAAGCAGC ACTCGTTGCC AATTAGGGA ATG GAC CGT TTG GGT TCC TTT AGC   173
                                Met Asp Arg Leu Gly Ser Phe Ser
                                                -45

AAT GAT CCC TCT GAT AAG CCA CCT TGC CGA GGC TGC TCC TCC TAC CTC   221
Asn Asp Pro Ser Asp Lys Pro Pro Cys Arg Gly Cys Ser Ser Tyr Leu
    -40                 -35                 -30

ATG GAG CCT TAT ATC AAG TGT GCT GAA TGT GGG CCA CCT CCT TTT TTC   269
Met Glu Pro Tyr Ile Lys Cys Ala Glu Cys Gly Pro Pro Pro Phe Phe
-25                 -20                 -15                 -10

CTC TGC TTG CAG TGT TTC ACT CGA GGC TTT GAG TAC AAG AAA CAT CAA   317
```

```
Leu Cys Leu Gln Cys Phe Thr Arg Gly Phe Glu Tyr Lys Lys His Gln
                -5                  1               5

AGC GAT CAT ACT TAT GAA ATA ATG GCA GGA TGT AGC CAA TCA AAT GTG     365
Ser Asp His Thr Tyr Glu Ile Met Ala Gly Cys Ser Gln Ser Asn Val
         10                  15                  20

CAC CAA GAC CAA GGA GGT CAG                                          386
His Gln Asp Gln Gly Gly Gln
     25                  30
```

(2) INFORMATION FOR SEQ ID NO: 144:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 389 base pairs
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: DOUBLE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
- (A) ORGANISM: Homo Sapiens
- (F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: 101..384
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 98
  region 1..284
  id R35114
  est

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: 90..365
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 91
  region 196..471
  id W65227
  est

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: 45..96
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 90
  region 152..203
  id W65227
  est

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: 90..349
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 93
  region 173..432
  id AA116536

est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 90..384
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 3..297
id AA065524
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 154..384
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..231
id R52412
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 57..317
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.6
seq GLLLLYMVYLTLV/EP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 144:

```
AAGAAAGCCG ACGGGGTGCT GTGGTCAGGA GGCGCGTGGG CGGCAGAATT TCGAGG ATG    59
                                                              Met

TCA GAT GTA AAT GTA TCT GCC CTC CCT ATA AAG AAA AAT TCT GGG CAT    107
Ser Asp Val Asn Val Ser Ala Leu Pro Ile Lys Lys Asn Ser Gly His
    -85                 -80                 -75

ATT TAT AAT AAG AAC ATA TCT CAG AAA GAT TGT GAT TGC CTT CAT GTC    155
Ile Tyr Asn Lys Asn Ile Ser Gln Lys Asp Cys Asp Cys Leu His Val
-70                 -65                 -60                 -55

GTG GAG CCC ATG CCT GTG CGG GGG CCT GAT GTA GAA GCA TAC TGT CTA    203
Val Glu Pro Met Pro Val Arg Gly Pro Asp Val Glu Ala Tyr Cys Leu
                -50                 -45                 -40

CGC TGT GAA TGC AAA TAT GAA GAA AGA AGC TCT GTC ACA ATC AAG GTT    251
Arg Cys Glu Cys Lys Tyr Glu Glu Arg Ser Ser Val Thr Ile Lys Val
            -35                 -30                 -25

ACC ATT ATA ATT TAT CTC TCC ATT TTG GGC CTT CTA CTT CTG TAC ATG    299
Thr Ile Ile Ile Tyr Leu Ser Ile Leu Gly Leu Leu Leu Leu Tyr Met
        -20                 -15                 -10

GTA TAT CTT ACT CTG GTT GAG CCC ATA CTG VAG AGG CGC CTC TTT GGA    347
Val Tyr Leu Thr Leu Val Glu Pro Ile Leu Xaa Arg Arg Leu Phe Gly
    -5                  1               5                  10

CAT GCA CAG TTG ATA CAG AGT GAT GAT RAT ATT GGG GGA TTG            389
His Ala Gln Leu Ile Gln Ser Asp Asp Xaa Ile Gly Gly Leu
                15                  20
```

(2) INFORMATION FOR SEQ ID NO: 145:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 228 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 45..150
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 25..130
id H30752
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 150..222
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 129..201
id H30752
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 93..222
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 91
region 76..205
id AA072341
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 73..156
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 1..84
id T31153
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 150..222
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 77..149
id T31153
est

```
(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 79..150
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 98
                              region 87..158
                              id W74452
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 152..222
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 100
                              region 158..228
                              id W74452
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 79..150
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 98
                              region 81..152
                              id H26405
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 150..222
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 97
                              region 151..223
                              id H26405
                              est

(ix) FEATURE:
      (A) NAME/KEY: sig_peptide
      (B) LOCATION: 85..222
      (C) IDENTIFICATION METHOD: Von Heijne matrix
      (D) OTHER INFORMATION:  score 3.5
                              seq ALSLSLSMAPPNP/GP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 145:

ACACCTCCCC GCCTTGTTGT CCAACTTCTC CCGGAGCAGC CGGAGAGCAG GCGTCGGGAC    60

GCAGCAAAGA GAGGAGAGGC CACC ATG GCG GAS TGC AGG AGG TGC AGA TCA      111
                           Met Ala Xaa Cys Arg Arg Cys Arg Ser
                               -45                 -40

CAG AGG AGA AGC CAC TGT TGC CAG GAC AGA CGC CTG AGG CGG CCA AGA     159
Gln Arg Arg Ser His Cys Cys Gln Asp Arg Arg Leu Arg Arg Pro Arg
        -35                 -30                 -25

CTC ACT CTG TGG AGA CAC CAT ACG GCT CTG TCA CTT TCA CTG TCT ATG     207
Leu Thr Leu Trp Arg His His Thr Ala Leu Ser Leu Ser Leu Ser Met
    -20                 -15                 -10

GCA CCC CCA AAC CCA GGC CCG                                         228
```

```
Ala Pro Pro Asn Pro Gly Pro
 -5                  1
```

(2) INFORMATION FOR SEQ ID NO: 146:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 376 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 246..375
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 9..138
id N56074
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 89..196
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq LALTALSVXRKXS/XX

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 146:

```
ACTGTCTGGC AGCCTGCAAG TCTCCATTCA GAAGCGGCTC CGTGCTGCCC AGCGATGGCG    60

CCCTGGCGGC GCGGAAGCCC GCGGCCAA ATG ACA CGA CTT GGG GGC AAA GGA      112
                               Met Thr Arg Leu Gly Gly Lys Gly
                                   -35                 -30

GGA CAA CAG TTC CCA CCA GGA CAA AAA ATA ATA TCC AAA GAT ATT TTG     160
Gly Gln Gln Phe Pro Pro Gly Gln Lys Ile Ile Ser Lys Asp Ile Leu
            -25                 -20                 -15

GCA CTA ACG GCG CTA TCT GTA GSA AGR AAG TTK AGC ART GKG RAM YKG     208
Ala Leu Thr Ala Leu Ser Val Xaa Arg Lys Xaa Ser Xaa Xaa Xaa Xaa
        -10                  -5                   1

TKN RAG ACT TCC AAG GAG ACT TBA GAC AAC CAA GAC AGT GTA AAG GAA     256
Xaa Xaa Thr Ser Lys Glu Thr Xaa Asp Asn Gln Asp Ser Val Lys Glu
  5                  10                  15                  20

AAC AGA GAA AAA GAC TTG TTA GAC ATT ATT AAG GGC ACG AAA GTT GAA     304
Asn Arg Glu Lys Asp Leu Leu Asp Ile Ile Lys Gly Thr Lys Val Glu
                 25                  30                  35

TTG AGC ACA GTA AAT GTA CAA ACA ACA AAG CCA CCC AAC AGA AGT TCA     352
```

```
Leu Ser Thr Val Asn Val Gln Thr Thr Lys Pro Pro Asn Arg Ser Ser
         40                   45                   50

CTT AAA AGC TAC AAC TGG CGG GCG                                  376
Leu Lys Ser Tyr Asn Trp Arg Ala
     55                   60
```

(2) INFORMATION FOR SEQ ID NO: 147:

```
    (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 265 base pairs
         (B) TYPE: NUCLEIC ACID
         (C) STRANDEDNESS: DOUBLE
         (D) TOPOLOGY: LINEAR

    (ii) MOLECULE TYPE: CDNA

    (vi) ORIGINAL SOURCE:
         (A) ORGANISM: Homo Sapiens
         (F) TISSUE TYPE: Substantia nigra

    (ix) FEATURE:
         (A) NAME/KEY: other
         (B) LOCATION: 123..265
         (C) IDENTIFICATION METHOD: blastn
         (D) OTHER INFORMATION:  identity 97
                                 region 187..329
                                 id N57089
                                 est

    (ix) FEATURE:
         (A) NAME/KEY: other
         (B) LOCATION: 47..127
         (C) IDENTIFICATION METHOD: blastn
         (D) OTHER INFORMATION:  identity 97
                                 region 110..190
                                 id N57089
                                 est

    (ix) FEATURE:
         (A) NAME/KEY: other
         (B) LOCATION: 3..218
         (C) IDENTIFICATION METHOD: blastn
         (D) OTHER INFORMATION:  identity 97
                                 region 2..217
                                 id AA136163
                                 est

    (ix) FEATURE:
         (A) NAME/KEY: other
         (B) LOCATION: 154..266
         (C) IDENTIFICATION METHOD: blastn
         (D) OTHER INFORMATION:  identity 98
                                 region 52..164
                                 id R22491
                                 est

    (ix) FEATURE:
         (A) NAME/KEY: other
```

```
        (B) LOCATION: 104..160
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 94
                                region 3..59
                                id R22491
                                est

   (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 47..150
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 99
                                region 1..104
                                id R29291
                                est

   (ix) FEATURE:
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: 194..253
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 12.4
                                seq ALLLGALLGTAWA/RR

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 147:

AAGGCGGTCG CCGGGACACC CCGTGTGTGG CAGGCGGCGA ASGCTCTGGA GAATCCCGGA    60

CAGCCCTGCT CCCTGCAGCC AGGTGTAGTT TCGGGAGCCA CTGGGGCCAA AGTGAGAGTC   120

CAGCGGTCTT CCAGCGCTTG GGCCACGGCG GCGGCCCTGG GAGCAGAGGT GGAGCGACCC   180

CATTACGCTA AAG ATG AAA GGC TGG GGT TGG CTG GCC CTG CTT CTG GGG      229
               Met Lys Gly Trp Gly Trp Leu Ala Leu Leu Leu Gly
               -20                 -15                 -10

GCC CTG CTG GGA ACC GCC TGG GCT CGG AGG AGC CGG                     265
Ala Leu Leu Gly Thr Ala Trp Ala Arg Arg Ser Arg
         -5                     1


(2) INFORMATION FOR SEQ ID NO: 148:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 222 base pairs
        (B) TYPE: NUCLEIC ACID
        (C) STRANDEDNESS: DOUBLE
        (D) TOPOLOGY: LINEAR

   (ii) MOLECULE TYPE: CDNA

   (vi) ORIGINAL SOURCE:
        (A) ORGANISM: Homo Sapiens
        (D) DEVELOPMENTAL STAGE: Fetal
        (F) TISSUE TYPE: brain

   (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 134..220
        (C) IDENTIFICATION METHOD: blastn
```

(D) OTHER INFORMATION: identity 91
region 14..100
id AA001733
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 88..135
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 11.4
seq LLCLLLLFGGGDP/RR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 148:

```
AAAGGTGCGC GTGCTCGCTG GTTCTAACCC TTCTGTTGGG CGTTTCTGCG GAGAGGCGGG   60

AGGCGCTGAG AGTCTGTGCG GAGGTCC ATG CAC AGA CTG CTT TGC CTG TTG TTG  114
                              Met His Arg Leu Leu Cys Leu Leu Leu
                                  -15                 -10

CTC TTC GGA GGC GGC GAT CCC CGA AGG CGA GCT GAA ATA CGG CTG CAG    162
Leu Phe Gly Gly Gly Asp Pro Arg Arg Arg Ala Glu Ile Arg Leu Gln
         -5                  1               5

GCT ACA ATT TGC AGC CGA CCA TTA AGG AAG ACG ACG AGC GGG AGA GGT    210
Ala Thr Ile Cys Ser Arg Pro Leu Arg Lys Thr Thr Ser Gly Arg Gly
 10                  15                  20                  25

GGC CCA CCC TGG                                                    222
Gly Pro Pro Trp
```

(2) INFORMATION FOR SEQ ID NO: 149:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 472 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 245..466
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 74..295
id R61190
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 181..258
(C) IDENTIFICATION METHOD: blastn

(D) OTHER INFORMATION: identity 94
region 11..88
id R61190
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 89..154
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 11.1
seq QLLALFFLPFCLC/QD

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 149:

```
AGCTCCAGTC CTGGCATCTG CCCGAGGAGA CCACGCTCCT GGAGCTCTGC TGTCTTCTCA    60

GGGAGACTCT GAGGCTCTGT TGAGAATC ATG CTT TGG AGG CAG CTC ATC TAT      112
                               Met Leu Trp Arg Gln Leu Ile Tyr
                                       -20                 -15

TGG CAA CTG CTG GCT TTG TTT TTC CTC CCT TTT TGC CTG TGT CAA GAT     160
Trp Gln Leu Leu Ala Leu Phe Phe Leu Pro Phe Cys Leu Cys Gln Asp
                -10                 -5                   1

GAA TAC ATG GAG TCT CCA CAA ACC GGA GGA CTA CCC CCA GAC TGC AGT     208
Glu Tyr Met Glu Ser Pro Gln Thr Gly Gly Leu Pro Pro Asp Cys Ser
         5                  10                  15

AAG TGT TGT CAT GGA GAC TAC AGC TTT CGA GGC TAC CAA GGC CCC CCT     256
Lys Cys Cys His Gly Asp Tyr Ser Phe Arg Gly Tyr Gln Gly Pro Pro
     20                  25                  30

GGG CCA CCG GGC CCT CCT GGC ATT CCA GGA AAC CAT GGA AAC AAT GGC     304
Gly Pro Pro Gly Pro Pro Gly Ile Pro Gly Asn His Gly Asn Asn Gly
 35                  40                  45                  50

AAC AAT GGA GCC ACT GGT CAT GAA GGA GCC AAA GGT GAG AAG GGC GAC     352
Asn Asn Gly Ala Thr Gly His Glu Gly Ala Lys Gly Glu Lys Gly Asp
                 55                  60                  65

AAA GGT GAC CTG GGG CCT CGA GGG GAG CGG GGG CAG CAT GGC CCC AAA     400
Lys Gly Asp Leu Gly Pro Arg Gly Glu Arg Gly Gln His Gly Pro Lys
             70                  75                  80

GGA GAG AAG GGC TAC CCG GGG ATT CCA CCA GAA CTT CAG ATT GCA TTC     448
Gly Glu Lys Gly Tyr Pro Gly Ile Pro Pro Glu Leu Gln Ile Ala Phe
         85                  90                  95

ATG GCT TCT CTG GMA CCC ACT TCA                                     472
Met Ala Ser Leu Xaa Pro Thr Ser
    100                 105
```

(2) INFORMATION FOR SEQ ID NO: 150:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 191 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 37..193
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 1..157
id T30099
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 53..193
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 1..141
id T35974
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 58..193
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 1..136
id T35248
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 58..193
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 1..136
id T32601
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 59..193
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 1..135
id T31945
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 81..131
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9.2
seq LLLLVAASAMVRS/XA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 150:

```
AAGTGGGGAG CAGCTCGCTC CTGGGCTTTG GGCTGGCTGC AGTCTGTCTG AGGGCGGCCG   60

AAGTGGCTGG CTCATTKAAG ATG AGG CTT CTA CTG CTT CTC CTA GTG GCG GCG  113
                      Met Arg Leu Leu Leu Leu Leu Leu Val Ala Ala
                                  -15                 -10

TCT GCG ATG GTC CGG AGC GAK GCC TCG GCC AAT CTG GGC GGC GTG CCC    161
Ser Ala Met Val Arg Ser Xaa Ala Ser Ala Asn Leu Gly Gly Val Pro
     -5                  1               5                  10

AGC AAG AGA TTA AAG ATG CAG TAC ACC ACG                            191
Ser Lys Arg Leu Lys Met Gln Tyr Thr Thr
                15                  20
```

(2) INFORMATION FOR SEQ ID NO: 151:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 561 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 136..210
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 177..251
id H23535
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 166..240
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9
seq ALLVLLGVAASLC/VR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 151:

```
AGGATGTACG GATGATTCAG TGGCTGGCAG GAAGCCCGCC CTGCCCGCCC GCCAGTGTCA   60

GTGGTGTTGG CATCAGCTTG GGCAGGTGTG CGGGCTCAGG ATGGGGCGGC CGTGGTGAGG  120

AACCCTGGAC TCTCAGCATC ACAAGAGGCA ACACCAGGAG CCAAC ATG AGC TCG GGG  177
                                                  Met Ser Ser Gly
                                                  -25

RCT GAA CTG CTG TGG CCC GGA GCA GCG CTG CTG GTG CTG TTG GGG GTG    225
Xaa Glu Leu Leu Trp Pro Gly Ala Ala Leu Leu Val Leu Leu Gly Val
```

```
    -20                 -15                 -10

GCA GCC AGT CTG TGT GTG CGC TGC TCA CGC CCA GGT GCA AAG AGG TCA      273
Ala Ala Ser Leu Cys Val Arg Cys Ser Arg Pro Gly Ala Lys Arg Ser
 -5                  1               5                   10

GAG AAA ATC TAC CAG CAG AGA AGT CTG CGT GAG GAC CAA CAG AGC TTT      321
Glu Lys Ile Tyr Gln Gln Arg Ser Leu Arg Glu Asp Gln Gln Ser Phe
             15                  20                  25

ACG GGG TCC CGG ACC TAC TCC TTG GTC GGG CAG GCA TGG CCA GGA CCC      369
Thr Gly Ser Arg Thr Tyr Ser Leu Val Gly Gln Ala Trp Pro Gly Pro
         30                  35                  40

CTG GCG GAC ATG GCA CCC ACA AGG AAG GAC AAG CTG TTG CAA TTC YAC      417
Leu Ala Asp Met Ala Pro Thr Arg Lys Asp Lys Leu Leu Gln Phe Xaa
     45                  50                  55

CCC AGC CTG GAG GMT CCA AGC ATC TTC CAG GKR MCA GAA MTT CAG CCA      465
Pro Ser Leu Glu Xaa Pro Ser Ile Phe Gln Xaa Xaa Glu Xaa Gln Pro
 60                  65                  70                  75

GTC TGT GTG TGC GCT GCT CAC GCC CAG GTG CAA ANN NVT CAG AGA AAA      513
Val Cys Val Cys Ala Ala His Ala Gln Val Gln Xaa Xaa Gln Arg Lys
                 80                  85                  90

TCT ACC AGC AGA GAA GTC TGC GTG AGG ACC AAC AGA GCT TTA CGG GGC      561
Ser Thr Ser Arg Glu Val Cys Val Arg Thr Asn Arg Ala Leu Arg Gly
             95                 100                 105
```

(2) INFORMATION FOR SEQ ID NO: 152:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 375 base pairs
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: DOUBLE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
- (A) ORGANISM: Homo Sapiens
- (D) DEVELOPMENTAL STAGE: Fetal
- (F) TISSUE TYPE: brain

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: 128..305
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 100
region 106..283
id H14437
est

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: 24..116
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 100

region 2..94
id H14437
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 317..376
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 36..95
id HSA86D041
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 1..234
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.8
seq SCLGLTLMPFASS/FP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 152:

```
ATG ACT AAG GAG ATT TTT TTT TTC ACA GTT GAG TTA GTT TGT GAA AAT      48
Met Thr Lys Glu Ile Phe Phe Phe Thr Val Glu Leu Val Cys Glu Asn
            -75                 -70                 -65

AAA GAA CTC TGT AGC TCA CCA AGG TGG AGA AAC GCA ATT CAG AAA AGT      96
Lys Glu Leu Cys Ser Ser Pro Arg Trp Arg Asn Ala Ile Gln Lys Ser
        -60                 -55                 -50

AAT TTC TCC AAG GTC ACT TCT TTT TTT ATG TCT TGC CAT CAC TTT AAA     144
Asn Phe Ser Lys Val Thr Ser Phe Phe Met Ser Cys His His Phe Lys
    -45                 -40                 -35

GGA CTA GCC CCA CTC CCC CAT GTG TAT ACA CAA GGA AAT TGC AGA CCA     192
Gly Leu Ala Pro Leu Pro His Val Tyr Thr Gln Gly Asn Cys Arg Pro
-30                 -25                 -20                 -15

ATT AGT TGT CTT GGC CTG ACT CTA ATG CCT TTT GCA AGT AGC TTT CCA     240
Ile Ser Cys Leu Gly Leu Thr Leu Met Pro Phe Ala Ser Ser Phe Pro
                -10                 -5                  1

GAA GTA AAA GTC CCA GTG ATG TAT TCC CAT AGA AAT ATT TTT CAG TTG     288
Glu Val Lys Val Pro Val Met Tyr Ser His Arg Asn Ile Phe Gln Leu
        5                   10                  15

TTT ATG TCG TTT ACT ACA AAA AAA AAG ATT CAG AGT GGA TGG AGT ACA     336
Phe Met Ser Phe Thr Thr Lys Lys Lys Ile Gln Ser Gly Trp Ser Thr
    20                  25                  30

ACT CTG AGT ATT TTT CTA GTC CGG AAT TTT TTA TTA ATA                 375
Thr Leu Ser Ile Phe Leu Val Arg Asn Phe Leu Leu Ile
35                  40                  45
```

(2) INFORMATION FOR SEQ ID NO: 153:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 339 base pairs

(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 209..332
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 39..162
id N42351
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 209..324
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 39..154
id N42357
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 209..284
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 25..100
id T57420
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 253..315
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.5
seq YFRALCLPRGAWG/FP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 153:

```
AAGGTGAGCT GAGGTACCTG GCTTTAGTGA AGACCCTGTG GGGGCTTCCT GGCTGCGGCT    60

TCGACRWGCC TGASTCCASA CGCCCCTAAG GTTTGATGAA AACCTGCTGG AGGTTTGGGA   120

CCTAACCGCA TCAAAGTCGC CTTTAGCGGT GCCTGGACCC AGTTCGCACG GGAGGAAGTA   180

GGAGGCAGAA TCCCCTTTGG GCCACAGAAA GCTCACCTGT TACTCGGCCT CCCAGAAAGA   240

TGGATAGGAG AA ATG ACT ACG GAT ATA GGG TGC CTC TAT TTC AGG GCC CTC   291
              Met Thr Thr Asp Ile Gly Cys Leu Tyr Phe Arg Ala Leu
                  -20                 -15                 -10

TGC CTC CCC CGG GGA GCC TGG GGC TTC CCT TCC CTC CAG ATT AAG GGG      339
Cys Leu Pro Arg Gly Ala Trp Gly Phe Pro Ser Leu Gln Ile Lys Gly
```

```
 -5                   1                  5

(2) INFORMATION FOR SEQ ID NO: 154:

      (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 112 base pairs
            (B) TYPE: NUCLEIC ACID
            (C) STRANDEDNESS: DOUBLE
            (D) TOPOLOGY: LINEAR

      (ii) MOLECULE TYPE: CDNA

      (vi) ORIGINAL SOURCE:
            (A) ORGANISM: Homo Sapiens
            (D) DEVELOPMENTAL STAGE: Fetal
            (F) TISSUE TYPE: brain

      (ix) FEATURE:
            (A) NAME/KEY: other
            (B) LOCATION: 10..109
            (C) IDENTIFICATION METHOD: blastn
            (D) OTHER INFORMATION:  identity 96
                                    region 1..100
                                    id W30713
                                    est

      (ix) FEATURE:
            (A) NAME/KEY: other
            (B) LOCATION: 24..109
            (C) IDENTIFICATION METHOD: blastn
            (D) OTHER INFORMATION:  identity 100
                                    region 1..86
                                    id HUM402E06B
                                    est

      (ix) FEATURE:
            (A) NAME/KEY: other
            (B) LOCATION: 44..109
            (C) IDENTIFICATION METHOD: blastn
            (D) OTHER INFORMATION:  identity 98
                                    region 30..95
                                    id H50196
                                    est

      (ix) FEATURE:
            (A) NAME/KEY: sig_peptide
            (B) LOCATION: 38..103
            (C) IDENTIFICATION METHOD: Von Heijne matrix
            (D) OTHER INFORMATION:  score 7.4
                                    seq LTCLFLFLNLRWS/RH

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 154:

AGTGTCTGCA CTTCGGCTGC TCTCGGGTTA GCACCCT ATG GTG CCT TCT CTT GTG    55
                                         Met Val Pro Ser Leu Val
                                                 -20

ATC CCT GAC CTA ACC TGT CTC TTC CTT TTC CTC AAC CTC AGG TGG AGC    103
```

```
Ile Pro Asp Leu Thr Cys Leu Phe Leu Phe Leu Asn Leu Arg Trp Ser
    -15                 -10                     -5

CGC CAC GTA                                                        112
Arg His Val
  1
```

(2) INFORMATION FOR SEQ ID NO: 155:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 191 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 89..189
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 71..171
id R11825
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 17..84
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 91
region 1..68
id R11825
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 98..189
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 83..174
id H08475
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 13..84
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 1..72
id H08475
est

(ix) FEATURE:
(A) NAME/KEY: other

(B) LOCATION: 30..84
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 49..103
id AA113990
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 150..189
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 163..202
id AA113990
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 89..141
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 67..119
id C14102
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 46..89
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 25..68
id C14102
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 160..189
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 141..170
id C14102
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 24..68
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7
seq LRLLKLAATSASA/RV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 155:

```
GATCCTGAGC TGACCGGGTA GCC ATG GCC TTG CGG CTC CTG AAG CTG GCA GCG    53
                          Met Ala Leu Arg Leu Leu Lys Leu Ala Ala
                          -15                 -10

ACG TCC GCG TCC GCC CGG GTC GTG RMG GCG GRM GCC CAG CGC GTG AGA     101
Thr Ser Ala Ser Ala Arg Val Val Xaa Ala Xaa Ala Gln Arg Val Arg
 -5                  1                   5                  10
```

```
GGA ATT CAT AGC AGT GTG CAG TGC AAG CTG CGC TAT GGA ATG TGG CAT      149
Gly Ile His Ser Ser Val Gln Cys Lys Leu Arg Tyr Gly Met Trp His
            15                  20                  25

TTC CTA CTT GGG GAT AAA GCA AGC AAA AGA CTG ACA GTA CAG              191
Phe Leu Leu Gly Asp Lys Ala Ser Lys Arg Leu Thr Val Gln
        30                  35                  40
```

(2) INFORMATION FOR SEQ ID NO: 156:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 160 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 24..161
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 1..138
id AA046377
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 75..161
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..87
id AA112337
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 33..62
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 333..362
id R72972
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 29..88
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.9
seq VLLFLYSVLLTKG/IE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 156:

```
AAAAGATGCT GTCTTGGACC AGTATTCA ATG TGG GGA AAT AAA TTT GGA GTA      52
                               Met Trp Gly Asn Lys Phe Gly Val
                               -20                 -15

TTG CTT TTT CTG TAT TCT GTA TTA CTG ACA AAG GGC ATT GAA AAC ATA    100
Leu Leu Phe Leu Tyr Ser Val Leu Leu Thr Lys Gly Ile Glu Asn Ile
        -10                 -5                  1

AAA AAC GAA ATT GAA GAT GCA AGT GAA CCC TTG ATA GAT CCT GTA TAT    148
Lys Asn Glu Ile Glu Asp Ala Ser Glu Pro Leu Ile Asp Pro Val Tyr
  5                   10                  15                  20

GGA CAT GGC ARC                                                    160
Gly His Gly Xaa
```

(2) INFORMATION FOR SEQ ID NO: 157:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 410 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(2..353)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..352
id N43180
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(2..332)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..331
id W56791
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 70..396
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 1..327
id T32797
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(123..396)

```
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 96
                                region 62..335
                                id T08229
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(109..396)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 97
                                region 62..349
                                id HSC27A061
                                est

    (ix) FEATURE:
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: 279..359
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 6.8
                                seq VFVCSSVLGQSWG/GF

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 157:

ATGTAGGACT TGTTCCTGTG GGCTTCAGTG ATGGGATAGT ACACTTCACT CAGAGGCATT   60

TGCATCTTTA AATAATTTCT TAAAAGCCTC TAAAGTGATC AGTGCCTTGA TGCCAACTAA  120

GGAAATTTGT TTAGCATTGA ATCTCTGAAG GCTCTATGAA AGGAATAGCA TGATGTGCTG  180

TTAGAATCAG ATGTTACTGC TAAAATTTAC ATGTTGTGAT GTAAATTGTG TAGAAAACCA  240

TTAAATCATT CAAAATAATA AACTATTTTT ATTAGAGA ATG TAT ACT TTT AGA AAG  296
                                          Met Tyr Thr Phe Arg Lys
                                                  -25

CTG TCT CCT TAT TTA AAT AAA ATA GTG TTT GTC TGT AGT TCA GTG TTG    344
Leu Ser Pro Tyr Leu Asn Lys Ile Val Phe Val Cys Ser Ser Val Leu
    -20                 -15                 -10

GGG CAA TCT TGG GGG GGA TTC TTC TCT AAT CTT TCA GAA ACT TTG TCT    392
Gly Gln Ser Trp Gly Gly Phe Phe Ser Asn Leu Ser Glu Thr Leu Ser
 -5                  1               5                  10

GCG ACA CTC TTT AAT GGA                                            410
Ala Thr Leu Phe Asn Gly
            15


(2) INFORMATION FOR SEQ ID NO: 158:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 188 base pairs
        (B) TYPE: NUCLEIC ACID
        (C) STRANDEDNESS: DOUBLE
        (D) TOPOLOGY: LINEAR

    (ii) MOLECULE TYPE: CDNA
```

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 3..186
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 3..186
id T31677
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 3..186
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 3..186
id HUM76142
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 39..186
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 1..148
id M85529
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 120..182
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.4
seq TFCLIFGLGAVWG/LV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 158:

```
AAAAATCCGA GGCAGCAGCA GGAGAGACAA ACGTTATTTT CCCGCTTGAT TCCAAGAACC   60

TCTTCGATAT TTATTTTTAT TTTTAAAGAG GGAGACGATG GACTGAGCTG ATCCGCACC   119

ATG GAG TCT CGG GTC TTA CTG AGA ACA TTC TGT TTG ATC TTC GGT CTC    167
Met Glu Ser Arg Val Leu Leu Arg Thr Phe Cys Leu Ile Phe Gly Leu
    -20                 -15                 -10

GGA GCA GTT TGG GGG CTG GTG                                        188
Gly Ala Val Trp Gly Leu Val
 -5                  1
```

(2) INFORMATION FOR SEQ ID NO: 159:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 218 base pairs
(B) TYPE: NUCLEIC ACID

```
        (C) STRANDEDNESS: DOUBLE
        (D) TOPOLOGY: LINEAR

   (ii) MOLECULE TYPE: CDNA

   (vi) ORIGINAL SOURCE:
        (A) ORGANISM: Homo Sapiens
        (D) DEVELOPMENTAL STAGE: Fetal
        (F) TISSUE TYPE: brain

   (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(37..103)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 97
                                region 2..68
                                id HSC07B041
                                est

   (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(2..38)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 97
                                region 68..104
                                id HSC07B041
                                est

   (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(1..38)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 97
                                region 60..97
                                id W23273
                                est

   (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(2..38)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 97
                                region 60..96
                                id W22817
                                est

   (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(2..37)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 91
                                region 54..89
                                id W23007
                                est

   (ix) FEATURE:
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: 33..143
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 6.4
                                seq ILIFLGFFLGLFH/QF
```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 159:

```
ACGGCCCGGC AGGAACAACT CATTCCGCTT AT ATG CTT GTT CTT AAA AAG CAC       53
                                    Met Leu Val Leu Lys Lys His
                                            -35

TCG GTA AAC ATT GCG GCC CAG ACG TGT TTT AAA TTC AAT TTT ATT TTC      101
Ser Val Asn Ile Ala Ala Gln Thr Cys Phe Lys Phe Asn Phe Ile Phe
-30                 -25                 -20                 -15

AGG ATC CTC ATC TTT CTT GGT TTC TTT CTG GGG CTT TTC CAT CAG TTC      149
Arg Ile Leu Ile Phe Leu Gly Phe Phe Leu Gly Leu Phe His Gln Phe
                -10                 -5                   1

CTC TTC CTC TTT CTC TTT GCT GGC AAT CTC AGC TCG TAC CTT TTG AAG      197
Leu Phe Leu Phe Leu Phe Ala Gly Asn Leu Ser Ser Tyr Leu Leu Lys
         5                  10                  15

CAG AGC AAA ATC CAA GCC AGG                                          218
Gln Ser Lys Ile Gln Ala Arg
     20                 25
```

(2) INFORMATION FOR SEQ ID NO: 160:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 314 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(110..315)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 186..391
id AA046808
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(8..105)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 394..491
id AA046808
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(110..315)

```
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 96
                                region 204..409
                                id AA156232
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(6..75)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 92
                                region 441..510
                                id AA156232
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 110..265
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 96
                                region 101..256
                                id AA147488
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 264..315
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 94
                                region 256..307
                                id AA147488
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 110..282
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 96
                                region 108..280
                                id AA157472
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 2..75
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 98
                                region 3..76
                                id AA157472
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 110..315
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 96
                                region 100..305
                                id AA046825
                                est

    (ix) FEATURE:
```

(A) NAME/KEY: other
(B) LOCATION: 8..75
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..68
id AA046825
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 165..260
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.3
seq TVVLCVGCSTVLC/QP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 160:

```
AGCTTGCAGC TTGCTAGCTG TGTGGGCTGG GAGGTCTGGT AGGGCTGAGC TTGCAAGAGG   60

RTCAACATGC CTTTGGCTAG AGAYKTAMTA CATRCGTMCY TGGRRWGWGG YABAAGAVMA  120

MACATAAAAA GAANCGCCTA GTACAAAGTC CAAATTCTTA CTTT ATG GAT GTA AAA  176
                                                 Met Asp Val Lys
                                                         -30

TGT CCA GGT TGC TAC AAG ATC ACC ACG GTT TTC AGC CAT GCT CAG ACA   224
Cys Pro Gly Cys Tyr Lys Ile Thr Thr Val Phe Ser His Ala Gln Thr
            -25                 -20                 -15

GTG GTT CTT TGT GTA GGT TGT TCA ACA GTG TTG TGC CAG CCT ACA GGA   272
Val Val Leu Cys Val Gly Cys Ser Thr Val Leu Cys Gln Pro Thr Gly
        -10                 -5                  1

GGA AAG GCC AGA CTC ACA GAA GGG TGT TCA TTT AGG MGA AAG           314
Gly Lys Ala Arg Leu Thr Glu Gly Cys Ser Phe Arg Arg Lys
  5                 10                  15
```

(2) INFORMATION FOR SEQ ID NO: 161:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 228 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 120..227
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 183..290

id T34489
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 80..119
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 144..183
id T34489
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 120..227
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 141..248
id HUM416B01B
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 80..119
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 102..141
id HUM416B01B
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 120..220
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 140..240
id HUM425D11B
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 80..119
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 101..140
id HUM425D11B
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 120..227
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 28..135
id AA147546
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 120..154
(C) IDENTIFICATION METHOD: blastn

```
        (D) OTHER INFORMATION:  identity 91
                                region 26..60
                                id AA103632
                                est

   (ix) FEATURE:
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: 163..210
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 6.2
                                seq ILSVLHALPAGIA/WS

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 161:

ACATTATCAA GAGAGAGAGA AGGTTAGGAT GAAAGAGATG AGCAGAGGCA AATCTTAAGT    60

GTTGACACCA TGATTGAAAA GGTTATTTGG AAATAGCGGT GCTTGCAGCG CYTGCGGAMC   120

RGTCGATTCC TGCGAGTGAA CTCGTCATGA GGGCGCAAGG CA ATG TGT ATC ATC      174
                                               Met Cys Ile Ile
                                                   -15

TTA AGT GTT TTA CAT GCT CTA CCT GCC GGA ATC GCC TGG TCC CGG GAG     222
Leu Ser Val Leu His Ala Leu Pro Ala Gly Ile Ala Trp Ser Arg Glu
        -10                 -5                   1

AAA GGG                                                             228
Lys Gly
  5


(2) INFORMATION FOR SEQ ID NO: 162:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 255 base pairs
        (B) TYPE: NUCLEIC ACID
        (C) STRANDEDNESS: DOUBLE
        (D) TOPOLOGY: LINEAR

   (ii) MOLECULE TYPE: CDNA

   (vi) ORIGINAL SOURCE:
        (A) ORGANISM: Homo Sapiens
        (F) TISSUE TYPE: Surrenals

   (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 58..255
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 98
                                region 176..373
                                id AA082793
                                est

   (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(97..255)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 98
```

```
                              region 352..510
                              id AA129762
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: complement(66..97)
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 100
                              region 511..542
                              id AA129762
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 89..214
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 98
                              region 11..136
                              id H56715
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 207..255
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 93
                              region 128..176
                              id H56715
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: complement(191..255)
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 98
                              region 355..419
                              id AA101128
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: complement(151..189)
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 97
                              region 423..461
                              id AA101128
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: complement(119..151)
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 90
                              region 462..494
                              id AA101128
                              est

(ix) FEATURE:
     (A) NAME/KEY: sig_peptide
     (B) LOCATION: 46..225
```

(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.9
seq TWLLLGALEPASE/RP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 162:

```
AGCAGCTGTT TCGGTAACTG CTTTGCCTCC CGGCTCCCGC AGWGG ATG CTG GTG GTG    57
                                                  Met Leu Val Val
                                                  -60

GAG GCT TCT TCC TCA GTG CGG CTG GCA AGT TCG GAG GTG ACT TCC TGG     105
Glu Ala Ser Ser Ser Val Arg Leu Ala Ser Ser Glu Val Thr Ser Trp
    -55                 -50                 -45

TCT ATC CTG GTG ACC CCC TCC GCT TCC ACG CCC ATT ATA TCG CTC AGT     153
Ser Ile Leu Val Thr Pro Ser Ala Ser Thr Pro Ile Ile Ser Leu Ser
-40                 -35                 -30                 -25

GCT GGG CCC CTG AGG ACA CCA TCC CAC TCC AAG ACC TGG TTG CTG CTG     201
Ala Gly Pro Leu Arg Thr Pro Ser His Ser Lys Thr Trp Leu Leu Leu
                -20                 -15                 -10

GGC GCC TTG GAA CCA GCG TCA GAA AGA CCC TGC TCC TCT GTT CTC CGC     249
Gly Ala Leu Glu Pro Ala Ser Glu Arg Pro Cys Ser Ser Val Leu Arg
            -5                   1               5

AGC CGG                                                             255
Ser Arg
    10
```

(2) INFORMATION FOR SEQ ID NO: 163:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 285 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 22..201
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 5..184
id W07565
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 200..282
(C) IDENTIFICATION METHOD: blastn

(D) OTHER INFORMATION: identity 100
region 184..266
id W07565
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 132..209
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 102..179
id R58430
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 208..282
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 177..251
id R58430
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 29..80
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 1..52
id R58430
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 82..132
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 53..103
id R58430
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 7..118
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..112
id W19152
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 175..282
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 109..216
id W19152
est

(ix) FEATURE:
(A) NAME/KEY: other

(B) LOCATION: 175..282
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 95..202
id AA095767
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 69..118
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 49..98
id AA095767
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 20..74
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 1..55
id AA095767
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 175..282
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 99..206
id T86637
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 22..118
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 6..102
id T86637
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 181..261
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.7
seq LSLQLIAFPTVSC/EI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 163:

```
CGCCTCCTCC GCTTTGGGAG CMCCGGGCTA MTCTTTCACA GCCCCTGTTG CCCTGTGATC    60

TGTAGGTCCT TGGGGACGCA CAGTTAAGAT GACAGGACAT CCTGGAAGCT GGGAAATGGC   120

TGAATGCTAT CCCAGTGAAT ATACGTGCCC TGTTTGTTGA ATCTACTCAT CCTTAAAGAT   180

ATG TAT TCA TTT CCT ACC ACC GTA GTG GAA GAG ATA CTA TCC CTA TCT     228
Met Tyr Ser Phe Pro Thr Thr Val Val Glu Glu Ile Leu Ser Leu Ser
```

```
        -25                 -20                 -15

TTA CAA CTG ATA GCA TTT CCA ACA GTA AGC TGT GAG ATT CTG CTT GAA     276
Leu Gln Leu Ile Ala Phe Pro Thr Val Ser Cys Glu Ile Leu Leu Glu
    -10                 -5                   1               5

ATC ACC AGG                                                         285
Ile Thr Arg
```

(2) INFORMATION FOR SEQ ID NO: 164:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 174 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 75..168
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 91
region 200..293
id N57841
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 115..162
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.6
seq LLPLRSLLALVRE/SR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 164:

```
AAAGCTTTAA CGAGCTTATA AATCACTTGG TAAATTTGAC CCCACTTTAT GTAATGTGAT    60

TCTGCAGGTT TGAAAAAGGT CCATAAATAG GTGTTTTAAA CAAGTTTCCT GTCA ATG     117
                                                            Met

CTG ATG TTG CTC CCC CTT CGC TCA TTA TTA GCA TTA GTT AGA GAA AGC     165
Leu Met Leu Leu Pro Leu Arg Ser Leu Leu Ala Leu Val Arg Glu Ser
-15                 -10                 -5                   1

AGG GCA CGG                                                         174
Arg Ala Arg
```

(2) INFORMATION FOR SEQ ID NO: 165:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 407 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 37..217
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 1..181
id AA057016
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 216..378
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 179..341
id AA057016
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 34..217
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..184
id AA133917
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 216..327
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 182..293
id AA133917
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 325..366
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 290..331
id AA133917
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 216..342
(C) IDENTIFICATION METHOD: blastn

(D) OTHER INFORMATION: identity 99
region 119..245
id R13065
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 114..217
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 18..121
id R13065
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 99..173
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.5
seq AQLFACLLRLGTQ/QV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 165:

```
AAGTCCAGGC CTTGAGACCC AGAAGGGAGC GAAGGTTTTT GCTGCGCCAA CGCAGTGAGC    60

CGAAGCTCCG CTCACGCCCG GCCTGATCCT GCCTGAAG ATG GTG CCA CTG GTG GCT   116
                                          Met Val Pro Leu Val Ala
                                          -25                 -20

GTG GTA TCA GGG CCC CGT GCC CAG CTC TTT GCC TGC CTG CTC AGG CTG     164
Val Val Ser Gly Pro Arg Ala Gln Leu Phe Ala Cys Leu Leu Arg Leu
            -15                 -10                 -5

GGC ACT CAG CAG GTC GGC CCC CTT CAG CTG CAC ACC GGG GCC AGC CAT     212
Gly Thr Gln Gln Val Gly Pro Leu Gln Leu His Thr Gly Ala Ser His
            1               5                   10

GCG GCC AGG AAC CAT TAT GAG GTG CTG GTG CTG GGT GGG GGC AGT GGC     260
Ala Ala Arg Asn His Tyr Glu Val Leu Val Leu Gly Gly Gly Ser Gly
    15                  20                  25

GGA ATC ACC ATG GCT GCC CGC ATG AAG AGG AAA GTG GGT GCA GAG AAT     308
Gly Ile Thr Met Ala Ala Arg Met Lys Arg Lys Val Gly Ala Glu Asn
30                  35                  40                  45

GTG GCC ATT GTT GAG CCC AGT GAG AGA CAT TTC TAC CAG CCA ATC TGG     356
Val Ala Ile Val Glu Pro Ser Glu Arg His Phe Tyr Gln Pro Ile Trp
                50                  55                  60

ACA CTG GTG GGT GCT GGT GCA ADC AAT TGT CCT CAT CTG GTC GTC CCA     404
Thr Leu Val Gly Ala Gly Ala Xaa Asn Cys Pro His Leu Val Val Pro
            65                  70                  75

CGG                                                                 407
Arg
```

(2) INFORMATION FOR SEQ ID NO: 166:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 184 base pairs
 (B) TYPE: NUCLEIC ACID
 (C) STRANDEDNESS: DOUBLE
 (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: Homo Sapiens
 (F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
 (A) NAME/KEY: other
 (B) LOCATION: 18..180
 (C) IDENTIFICATION METHOD: blastn
 (D) OTHER INFORMATION: identity 93
 region 18..180
 id AA110680
 est

(ix) FEATURE:
 (A) NAME/KEY: other
 (B) LOCATION: 40..180
 (C) IDENTIFICATION METHOD: blastn
 (D) OTHER INFORMATION: identity 95
 region 64..204
 id W30470
 est

(ix) FEATURE:
 (A) NAME/KEY: other
 (B) LOCATION: 91..180
 (C) IDENTIFICATION METHOD: blastn
 (D) OTHER INFORMATION: identity 94
 region 46..135
 id HSCOCC021
 est

(ix) FEATURE:
 (A) NAME/KEY: other
 (B) LOCATION: 91..180
 (C) IDENTIFICATION METHOD: blastn
 (D) OTHER INFORMATION: identity 100
 region 23..112
 id HUMHG5097
 est

(ix) FEATURE:
 (A) NAME/KEY: other
 (B) LOCATION: 80..180
 (C) IDENTIFICATION METHOD: blastn
 (D) OTHER INFORMATION: identity 94
 region 16..116
 id AA089700
 est

(ix) FEATURE:
 (A) NAME/KEY: sig_peptide
 (B) LOCATION: 110..169
 (C) IDENTIFICATION METHOD: Von Heijne matrix

```
        (D) OTHER INFORMATION:  score 5.4
                                seq AFVIACVLSLIST/IY

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 166:

AGAGCTAGAG GGTGAAGCTG GCGGAGCAGG AGGATGGGCG TATGCAGGTG ATAGACTAGA    60

GAACAAGACC TCTGTCTCCG TAGCATCCTG GAGCAGTCTG AATGCCAGA ATG GAT AAC   118
                                                      Met Asp Asn
                                                      -20

CGT TTT GCT ACA GCA TTT GTA ATT GCT TGT GTG CTT AGC CTC ATT TCC     166
Arg Phe Ala Thr Ala Phe Val Ile Ala Cys Val Leu Ser Leu Ile Ser
        -15                 -10                 -5

ACC ATC TAC ATG GCC CGG                                             184
Thr Ile Tyr Met Ala Arg
    1               5


(2) INFORMATION FOR SEQ ID NO: 167:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 371 base pairs
        (B) TYPE: NUCLEIC ACID
        (C) STRANDEDNESS: DOUBLE
        (D) TOPOLOGY: LINEAR

    (ii) MOLECULE TYPE: CDNA

    (vi) ORIGINAL SOURCE:
        (A) ORGANISM: Homo Sapiens
        (F) TISSUE TYPE: Substantia nigra

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(1..223)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 99
                                region 277..499
                                id AA059399
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(236..330)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 100
                                region 170..264
                                id AA059399
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(332..371)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 95
                                region 130..169
                                id AA059399
```

est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(6..143)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 357..494
id N35134
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(236..371)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 128..263
id N35134
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(141..223)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 276..358
id N35134
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(44..223)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 223..402
id H49423
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(236..371)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 75..210
id H49423
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 14..190
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 1..177
id H43799
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 236..342
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98

region 224..330
id H43799
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 338..369
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 327..358
id H43799
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(85..223)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 223..361
id H27862
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(266..346)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 101..181
id H27862
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(236..269)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 177..210
id H27862
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 12..137
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.4
seq WTLLLTSLDGHLL/EP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 167:

```
AGGATGACCT G ATG CCT GAG TAC TGT GGC AAT GAG GTG ACT CCA ACC GAG      50
             Met Pro Glu Tyr Cys Gly Asn Glu Val Thr Pro Thr Glu
                     -40                 -35                 -30

GCT GCC CAA GCG CCA GAG GTG ACC TAT GAG GCA GAA GAG GGC TCC TTG       98
Ala Ala Gln Ala Pro Glu Val Thr Tyr Glu Ala Glu Glu Gly Ser Leu
                -25                 -20                 -15

TGG ACG TTG CTA CTC ACT AGC TTG GAT GGG CAC CTG CTG GAG CCA GAT      146
Trp Thr Leu Leu Leu Thr Ser Leu Asp Gly His Leu Leu Glu Pro Asp
            -10                  -5                  1
```

```
GCT GAG TAC CTC CAC TGG CTG CTA ACC AAC ATC CCG GGT AAC CGG GTG     194
Ala Glu Tyr Leu His Trp Leu Leu Thr Asn Ile Pro Gly Asn Arg Val
      5                  10                      15

GCT GAA GGA CAG GTG ACG TGT CCC TAC CTC CCC CCC TTC CCT GCC CGA     242
Ala Glu Gly Gln Val Thr Cys Pro Tyr Leu Pro Pro Phe Pro Ala Arg
 20                  25                      30                  35

GGC TCC GGC ATC CAC CGT CTT GCC TTC CTG CTC TTC AAG CAG GAC CAG     290
Gly Ser Gly Ile His Arg Leu Ala Phe Leu Leu Phe Lys Gln Asp Gln
                 40                  45                      50

CCG ATT GAC TTC TCT GAG GAC GCA CGC CCC TCA CCC TGC TAT CAG CTG     338
Pro Ile Asp Phe Ser Glu Asp Ala Arg Pro Ser Pro Cys Tyr Gln Leu
             55                  60                      65

GYS CAG CGG ACC TTC CGC ACT TTT GAT TTC TAC                         371
Xaa Gln Arg Thr Phe Arg Thr Phe Asp Phe Tyr
         70                  75
```

(2) INFORMATION FOR SEQ ID NO: 168:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 118 base pairs
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: DOUBLE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
- (A) ORGANISM: Homo Sapiens
- (F) TISSUE TYPE: Surrenals

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: 51..114
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 98
region 112..175
id AA029014
est

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: 20..55
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 100
region 1..36
id AA029014
est

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: 5..55
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 98
region 1..51

id AA029165
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 70..114
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 66..110
id AA029165
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 29..58
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 8..37
id W79853
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 23..67
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.4
seq RVLCAPAAGAVRA/LR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 168:

```
AAGTCGAGGA GTCAAGGCAG CA ATG AAT CGT GTC TTG TGT GCC CCG GCG GCC     52
                         Met Asn Arg Val Leu Cys Ala Pro Ala Ala
                         -15                 -10

GGG GCC GTC CGG GCG CTG AGG CTC ATA GGC TGG GCT TCC CGA AGC CTT     100
Gly Ala Val Arg Ala Leu Arg Leu Ile Gly Trp Ala Ser Arg Ser Leu
 -5                 1               5                  10

CAT CCG TTG CCC GGA AAG                                             118
His Pro Leu Pro Gly Lys
            15
```

(2) INFORMATION FOR SEQ ID NO: 169:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 183 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other

```
        (B) LOCATION: 123..181
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 96
                                region 133..191
                                id W02973
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 4..55
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 100
                                region 13..64
                                id W02973
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 56..102
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 100
                                region 66..112
                                id W02973
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 55..181
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 100
                                region 38..164
                                id T97581
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 19..56
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 100
                                region 1..38
                                id T97581
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 21..181
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 91
                                region 58..218
                                id W70849
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 21..181
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 91
                                region 47..207
                                id AA049525
                                est
```

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 22..181
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 91
region 4..163
id H32732
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 49..117
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.3
seq MLALLLTAALIFF/AI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 169:

```
ACTTTCTCCG CTGGCAACGG CGCCGCTCCC CGCTCCTCCT CCCCAGCC ATG GCG TTC      57
                                                     Met Ala Phe

ACG TTC GCG GCC TTC TGC TAC ATG CTG GCG CTG CTG CTC ACT GCC GCG      105
Thr Phe Ala Ala Phe Cys Tyr Met Leu Ala Leu Leu Leu Thr Ala Ala
-20                 -15                 -10                  -5

CTC ATC TTC TTC GCC ATT TGG CAC ATT ATA GCA TTT GAT GAG CTG AAG      153
Leu Ile Phe Phe Ala Ile Trp His Ile Ile Ala Phe Asp Glu Leu Lys
                  1               5                  10

ACT GAT TAC AAG AAT CCT ATA GAC CAG TTG                              183
Thr Asp Tyr Lys Asn Pro Ile Asp Gln Leu
        15                  20
```

(2) INFORMATION FOR SEQ ID NO: 170:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 169 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 107..143
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 125..161
id N86955
est

(ix) FEATURE:

(A) NAME/KEY: sig_peptide
(B) LOCATION: 89..139
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.3
seq XEXLLAFHHDCEA/SP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 170:

```
AATAGGTAAT GAGTCTTATG AGATCTGTTG GTTTTAAAAA CAGGAGTTTC TCTGCGCAAS   60

CTCTGTCTTT TTTTGCCTGC TGGCATCC ATG CRA RRC RWG AST GAG YTC CTC      112
                               Met Xaa Xaa Xaa Xaa Glu Xaa Leu
                                       -15                 -10

CTT GCC TTC CAC CAT GAT TGT GAG GCT TCC CCA GCC ACG TGG AAC TTA     160
Leu Ala Phe His His Asp Cys Glu Ala Ser Pro Ala Thr Trp Asn Leu
                -5                  1               5

AGT CCA AGG                                                         169
Ser Pro Arg
         10
```

(2) INFORMATION FOR SEQ ID NO: 171:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 240 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(15..236)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 314..535
id AA194996
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 13..120
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.2
seq PLRLLNLLILIEG/SV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 171:

```
TTATATAGAG CC ATG GGG CCT TAC AAC GTG GCA GTG CCT TCA GAT GTA TCT    51
              Met Gly Pro Tyr Asn Val Ala Val Pro Ser Asp Val Ser
```

```
                 -35                      -30                      -25

CAT GCC CGC TTT TAT TTC TTA TTT CAT CGA CCA TTA AGG CTG TTA AAT      99
His Ala Arg Phe Tyr Phe Leu Phe His Arg Pro Leu Arg Leu Leu Asn
            -20                     -15                     -10

CTG CTC ATC CTT ATT GAG GGC AGT GTC GTC TTC TAT CAG CTC TAT TCC     147
Leu Leu Ile Leu Ile Glu Gly Ser Val Val Phe Tyr Gln Leu Tyr Ser
        -5                       1               5

TTG CTG CGG TCG GAG AAG TGG AAC CAC ACA CTT TCC ATG GCT CTC ATC     195
Leu Leu Arg Ser Glu Lys Trp Asn His Thr Leu Ser Met Ala Leu Ile
10                  15                  20                  25

CTC TTC TGC AAC TAC TAT GTT TTA TTT AAA CTT CTC CGG GAT CAG         240
Leu Phe Cys Asn Tyr Tyr Val Leu Phe Lys Leu Leu Arg Asp Gln
                30                  35                  40
```

(2) INFORMATION FOR SEQ ID NO: 172:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 207 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 31..207
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..177
id W44975
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 45..207
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..163
id N43016
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 47..207
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..161
id R10507
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 49..207
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..159
id H70117
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 48..207
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..160
id H61870
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 67..135
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.2
seq IGVGLYLLASAAA/FY

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 172:

```
AGCGGCGGCA TCCGGGACGG CGGGCGGGCT GGCCACCACG GGACAGGAAG GCACAGAGCA   60

TGGAGA ATG ATG AAC TTC CGT CAG CGG ATG GGA TGG ATT GGA GTG GGA     108
       Met Met Asn Phe Arg Gln Arg Met Gly Trp Ile Gly Val Gly
                   -20                 -15                 -10

TTG TAT CTG TTA GCC AGT GCA GCA GCA TTT TAC TAT GTT TTT GAA ATC    156
Leu Tyr Leu Leu Ala Ser Ala Ala Ala Phe Tyr Tyr Val Phe Glu Ile
                -5                   1               5

AGT GAG ACT TAC AAC AGG CTG GCC TTG GAA CAC ATT CAA CAG CAC CHT    204
Ser Glu Thr Tyr Asn Arg Leu Ala Leu Glu His Ile Gln Gln His Xaa
        10                  15                  20

GGG                                                                207
Gly
```

(2) INFORMATION FOR SEQ ID NO: 173:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 487 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

```
(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 71..252
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 97
                             region 149..330
                             id R18686
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 252..304
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 100
                             region 331..383
                             id R18686
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 23..73
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 100
                             region 1..51
                             id R18686
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 186..375
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 99
                             region 1..190
                             id R54039
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 376..489
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 94
                             region 190..303
                             id R54039
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 94..303
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 98
                             region 70..279
                             id HSC2LA121
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 26..97
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 93
                             region 1..72
```

id HSC2LA121
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 236..375
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 1..140
id N91698
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 376..458
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 140..222
id N91698
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 458..489
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 223..254
id N91698
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 98..451
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.1
seq AFXVVCWLGPCEA/MH

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 173:

```
AAGCGCCCGC VGSSCGCGTC CCCGGCCCAA CCATGGCGTC CTCCGCGGCC GGCTGCGTGG    60

TGATCGTTGG CAGTGGAGTC ATTGGGCGAR STGGGCC ATG CTG TTT GCC AGT GGA   115
                                         Met Leu Phe Ala Ser Gly
                                                     -115

GGC TTC CAK GTG AAA CTC TAT GAC ATT GAG CAA CAG CAG ATA AGG AAC   163
Gly Phe Xaa Val Lys Leu Tyr Asp Ile Glu Gln Gln Gln Ile Arg Asn
        -110                -105                -100

GCC CTG GAA AAC ATC AGA AAG GAG ATG AAG TTG CTG GAG CAG GCA GGT   211
Ala Leu Glu Asn Ile Arg Lys Glu Met Lys Leu Leu Glu Gln Ala Gly
    -95                 -90                 -85

TCT CTG AAA GGC TCC CTG AGT GTG GAA GAG CAG CTG TCA CTC ATC AGT   259
Ser Leu Lys Gly Ser Leu Ser Val Glu Glu Gln Leu Ser Leu Ile Ser
-80                 -75                 -70                 -65

GGT TGT CCC AAT ATC CAA GAA GCA GTA GAG GGT GCC ATG CAC ATT CAG   307
Gly Cys Pro Asn Ile Gln Glu Ala Val Glu Gly Ala Met His Ile Gln
                -60                 -55                 -50
```

```
GAA TGT GTT CCA GAA GAT CTA GAA CTG AAG AAG AAG ATT TTT GCT CAG     355
Glu Cys Val Pro Glu Asp Leu Glu Leu Lys Lys Lys Ile Phe Ala Gln
            -45                 -40                 -35

TTA GAT TCC ATC ATT GAT GAA TCG AGT GAT CTT AAG CGT TTC CAM TTC     403
Leu Asp Ser Ile Ile Asp Glu Ser Ser Asp Leu Lys Arg Phe Xaa Phe
        -30                 -25                 -20

TTG TCT CAT GCC TTC VAA GTT GTT TGC TGG CTT GGT CCA TGT GAA GCA     451
Leu Ser His Ala Phe Xaa Val Val Cys Trp Leu Gly Pro Cys Glu Ala
    -15                 -10                 -5

ATG CAT CGT GGM TCA TCC TGT GAA TCC GCC ATA CTA                     487
Met His Arg Gly Ser Ser Cys Glu Ser Ala Ile Leu
  1               5                  10
```

(2) INFORMATION FOR SEQ ID NO: 174:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 140 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(47..141)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 236..330
id R62451
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(78..141)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 224..287
id C14686
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 78..141
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..64
id C15352
est

(ix) FEATURE:

(A) NAME/KEY: other
(B) LOCATION: 78..141
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..64
id C15019
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(78..141)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 225..288
id C14869
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 39..104
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.1
seq LCSLPLSPSAVCP/AP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 174:

```
ACACACAAAA ATTTTCCTTT GTTTGCGGGG GGCTGGGG ATG CAG TGT TTT TTG GGG    56
                                          Met Gln Cys Phe Leu Gly
                                                  -20

GGT CTT GGT TTA TGC TCC CTG CCC TTG AGC CCC TCA GCC GTT TGC CCT     104
Gly Leu Gly Leu Cys Ser Leu Pro Leu Ser Pro Ser Ala Val Cys Pro
    -15                 -10                     -5

GCC CCC ACC TCG GCT CCA TGG TGG GAG GGG GCT CTG                     140
Ala Pro Thr Ser Ala Pro Trp Trp Glu Gly Ala Leu
  1               5                  10
```

(2) INFORMATION FOR SEQ ID NO: 175:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 363 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 37..288
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 4..255

id R13070
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 287..353
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 255..321
id R13070
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 175..213
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.9
seq MSSFLLSFSQSLS/NV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 175:

```
ACTCTTTACT CCCCTGTGAG TGATTCACTG CCTTGTCATT ATTACGATAG ATGTGTTTGT   60

ATTGTKTTTT TTCTGATGAT ACTGATGTTG ATGAATTTTT AATTTTATTT GATGTGGTAG  120

AGTTGGGAGG TTTCAGGGTT TTTTCCCCTC TTTTACTTTC CATTGAGGAA GGGA ATG    177
                                                            Met

AGC TCC TTT CTC CTC TCC TTC AGC CAA TCA TTA TCA AAT GTT CCT TCA    225
Ser Ser Phe Leu Leu Ser Phe Ser Gln Ser Leu Ser Asn Val Pro Ser
        -10                 -5                   1

GCC CTG CAG TKG CCC CAA ATA ACC TTT TTT CAG CAT CCT CTG TCC TCA    273
Ala Leu Gln Xaa Pro Gln Ile Thr Phe Phe Gln His Pro Leu Ser Ser
  5                  10                  15                  20

GTC ATG CCA GTC TGG ACA TGC TCT GTT GTG CCC TGT GAC AAA ACT GST    321
Val Met Pro Val Trp Thr Cys Ser Val Val Pro Cys Asp Lys Thr Xaa
                 25                  30                  35

CAG TAT TCC TAT TGC TTT TAC TGT GTT TTA GGT ACT GTG AAG            363
Gln Tyr Ser Tyr Cys Phe Tyr Cys Val Leu Gly Thr Val Lys
             40                  45                  50
```

(2) INFORMATION FOR SEQ ID NO: 176:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 379 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

```
(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: complement(216..354)
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 98
                              region 82..220
                              id W28597
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: complement(351..381)
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 93
                              region 54..84
                              id W28597
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 216..347
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 99
                              region 121..252
                              id T90405
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 249..310
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 90
                              region 141..202
                              id AA085837
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 293..338
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 95
                              region 186..231
                              id AA085837
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 216..248
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 96
                              region 107..139
                              id AA085837
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 351..381
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 93
                              region 246..276
                              id AA085837
```

est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 152..190
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.9
seq MLTASLAFQLVDG/VS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 176:

```
AGAATCTGCT CGCTCTTCTG GAGCTGAGAC ACTCATCTTC TCCTGCCCTG GCACATCAAA   60

ACTCCAGGTT CTCTGATCTT TGAACACTGG GACTTAATCC AGCATCCCCT TACTGCCTGA  120

AGTTCTCAGG ACTTTGGACT TGGACTGAAC G ATG CTA ACA GCT TCC CTG GCT     172
                                   Met Leu Thr Ala Ser Leu Ala
                                                -10

TTC CAG CTT GTA GAT GGC GTA TCA TGG AAT TTC TCA GTC TCC AAA ATG    220
Phe Gln Leu Val Asp Gly Val Ser Trp Asn Phe Ser Val Ser Lys Met
     -5                  1                   5                  10

CTG GCA TCA CCA TCT ACA TCA GGT CAG CTG TCT CAG TTT GGG GCA AGT    268
Leu Ala Ser Pro Ser Thr Ser Gly Gln Leu Ser Gln Phe Gly Ala Ser
                15                  20                  25

TTA TAC GGG CAA CAA AGT GCA CTA GGC CTT CCA ATG AGG GGG ATG AGC    316
Leu Tyr Gly Gln Gln Ser Ala Leu Gly Leu Pro Met Arg Gly Met Ser
            30                  35                  40

AAC AAT ACC CCT CAG TTA AAT CGC AGC TTA TCA CAA GSA CTC AGT TAC    364
Asn Asn Thr Pro Gln Leu Asn Arg Ser Leu Ser Gln Xaa Leu Ser Tyr
        45                  50                  55

CGA GCC ACG TCA CGC                                                379
Arg Ala Thr Ser Arg
     60
```

(2) INFORMATION FOR SEQ ID NO: 177:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 198 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(2..132)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99

region 105..235
id T58540
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 118..192
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.9
seq LSAFNFLVCLSLG/RG

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 177:

```
ATATGATGCT GGCAAGACAC CCAGAAAGGC TATTTTCAGA TGAAATCGAT ATTAGAAGCT    60

ATATTAGCTG AAACAACTCC TTTTACTGCG TAGAACCTAT ATCGAGAGTG TGTGTAT      117

ATG TAT TAW AGG AGG GAG CTC TCA ATT TTA TGT ATT CTT TCT GCC TTT     165
Met Tyr Xaa Arg Arg Glu Leu Ser Ile Leu Cys Ile Leu Ser Ala Phe
-25                 -20                 -15                 -10

AAT TTT CTT GTT TGT TTG AGC TTA GGG AGA GGG                         198
Asn Phe Leu Val Cys Leu Ser Leu Gly Arg Gly
                 -5                  1
```

(2) INFORMATION FOR SEQ ID NO: 178:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 292 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 56..93
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 213..250
id N58549
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 206..271
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.8
seq IVFGVSWVMLVYS/AS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 178:

```
ATGWAGGCTG TTGTTCAGAT TCCTTTGTCC CATGGGGTGT TCCCTTATTA TAGTACTCTC   60

GCCCTTTTCC TATGGATGTG GCTTCCTGAG AGCGGMSTGA TASTGATTGT KATCTTTCTT  120

KTGGATCTAS CCACCCAGCT ATTCTACCAK GCTCTGGGCT GGTACTGGCG GTTGTCTGCA  180

CAGAGTCTTG TGACCTGAAC CATCT ATG GGT CTC TCW GCC ATG GAT ACC AGC    232
                            Met Gly Leu Ser Ala Met Asp Thr Ser
                                    -20                 -15

ATA GTA TTT GGG GTG TCC TGG GTC ATG CTG GTG TAC TCT GCT TCC TTC    280
Ile Val Phe Gly Val Ser Trp Val Met Leu Val Tyr Ser Ala Ser Phe
            -10                 -5                   1

AGG AGG TGT GRN                                                    292
Arg Arg Cys Xaa
    5
```

(2) INFORMATION FOR SEQ ID NO: 179:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 307 base pairs
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: DOUBLE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
- (A) ORGANISM: Homo Sapiens
- (F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: complement(80..196)
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 100
  region 68..184
  id R85971
  est

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: complement(31..86)
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 96
  region 179..234
  id R85971
  est

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: complement(80..196)
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 100
  region 68..184
  id R85965
  est

```
(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: complement(31..86)
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 96
                              region 179..234
                              id R85965
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 205..305
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 95
                              region 128..228
                              id N33907
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 205..305
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 95
                              region 99..199
                              id T19230
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 205..305
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 95
                              region 96..196
                              id T31783
                              est

(ix) FEATURE:
      (A) NAME/KEY: sig_peptide
      (B) LOCATION: 101..295
      (C) IDENTIFICATION METHOD: Von Heijne matrix
      (D) OTHER INFORMATION:  score 4.8
                              seq IILFSAIVGFIYG/YV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 179:

AAAAAAGGTC AAGGTAGCCT CCTCCAAACT CCACGTTGAG CTGCACCTCC CCTGAGCTCC   60

CACATCTGGG CCTCCCTCTC CCAGGGCTAA TCCAGCTGTC ATG TAT TTC TGG CGT   115
                                            Met Tyr Phe Trp Arg
                                            -65

GAT GTG GCT GTC TCC CTT GAC ACG CTC TGG GCT CTT CCA AGG CAA CAG   163
Asp Val Ala Val Ser Leu Asp Thr Leu Trp Ala Leu Pro Arg Gln Gln
-60                 -55                 -50                 -45

CCT GGT CTT GGT AAC AAC CGT GTC CTC GGT CTG CTG TCT GGC ACA AAC   211
Pro Gly Leu Gly Asn Asn Arg Val Leu Gly Leu Leu Ser Gly Thr Asn
                -40                 -35                 -30

AAG GAT TAC AAG GGC CAG AAG CTA GCT GAA CAG ATG TTT CAG GGA ATT   259
```

```
Lys Asp Tyr Lys Gly Gln Lys Leu Ala Glu Gln Met Phe Gln Gly Ile
            -25                 -20                 -15

ATT CTT TTT TCT GCA ATA GTT GGA TTT ATC TAC GGG TAC GTG GCC GCG      307
Ile Leu Phe Ser Ala Ile Val Gly Phe Ile Tyr Gly Tyr Val Ala Ala
        -10                  -5                   1
```

(2) INFORMATION FOR SEQ ID NO: 180:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 260 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 24..261
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 1..238
id HSC29G021
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 99..176
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.7
seq LLICXLXIGTATP/VR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 180:

```
ATTACGCAGA GAGAAAGTTA CGAGGTTCGT GGCCGCGGTT TCCCCAGGCA GCTGGCGCTG    60

GAGGCTTCGG CGTCACGTGC TGGTCTGGRT TTTTCTCG ATG CAC TGG GGA AAG CGG   116
                                          Met His Trp Gly Lys Arg
                                                  -25

TGG RCT CTT RTC GWR GGA GGG CTC TTG ATC TGT GRT TTA TRG ATA GGC      164
Trp Xaa Leu Xaa Xaa Gly Gly Leu Leu Ile Cys Xaa Leu Xaa Ile Gly
-20                 -15                 -10                  -5

ACA GCT ACT CCC GTT CGG GRA CCC AAC GGC AGA CAG GTS CTA GTG CCC      212
Thr Ala Thr Pro Val Arg Xaa Pro Asn Gly Arg Gln Val Leu Val Pro
                  1                   5                  10

RTC GGW TAC CCG CGG CCG GGA CTC GGA GCT GTG GGG TGT GGG GAG GCG      260
Xaa Gly Tyr Pro Arg Pro Gly Leu Gly Ala Val Gly Cys Gly Glu Ala
        15                  20                  25
```

(2) INFORMATION FOR SEQ ID NO: 181:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 416 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 158..364
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 91
region 153..359
id N25870
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 358..416
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 91
region 355..413
id N25870
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 148..332
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 91
region 134..318
id AA045920
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 332..395
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 319..382
id AA045920
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 205..416
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 1..212
id AA150024
est

(ix) FEATURE:

```
        (A) NAME/KEY: other
        (B) LOCATION: 148..315
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 90
                                region 133..300
                                id H99323
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 352..403
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 92
                                region 337..388
                                id H99323
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(202..384)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 93
                                region 215..397
                                id R83259
                                est

    (ix) FEATURE:
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: 309..395
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 4.6
                                seq ALGLXTCLSVLFG/YA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 181:

AGCASGGAAC AGCGGGTGCG GACATTACGG CGGAAGGCTC TSGAGGAAGC AGAAGTGAAG    60

GACCTCGCAS TCCTGGGACG GTGGGGCMCA GASAGAGAAA GGGAGCCCCG GGCGCGGCGC   120

GGTGAGGATG CGAGCAGAGG AAGGACACGC GGCGCCGSAA AATATTTACA CCAGCAGCTC   180

CAGTTCATAC CMMTAAAGAM SATCCTGCTA CCCAMACTAM TTTGGGATTT ATCCMYGCAT   240

TTGTCGCTGC CATATCAGTT ATTATTGTAT CTGAATTGGG TGATAAGACA TTTTTTATAG   300

CAGCCATC ATG GCA MCG CGC TAT AAC CGC CTG ACC GTG CTG GCT GGT GCM     350
         Met Ala Xaa Arg Tyr Asn Arg Leu Thr Val Leu Ala Gly Ala
                         -25                 -20

MTG CTT GCC TTG GGA CTA MTG ACA TGC YTG TCA GTT TTG TTT GGC TAT      398
Xaa Leu Ala Leu Gly Leu Xaa Thr Cys Leu Ser Val Leu Phe Gly Tyr
-15                 -10                 -5                    1

GCA CCA CAG TCA TCC CCA                                              416
Ala Pro Gln Ser Ser Pro
              5


(2) INFORMATION FOR SEQ ID NO: 182:
```

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 247 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 55..152
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 188..285
id N94950
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 54..152
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 122..220
id T09182
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 170..233
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 296..359
id N71787
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(54..142)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 199..287
id T64627
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 143..229
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.5
seq FMTCILCRPPISS/CV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 182:

ACAGATCTGA GGGTTCCTCG CCTTTTTAGA CCATATAGGG TAGCCTTCTG ATGTGCCATG 60

```
GTATTTGTAA ACTGTCATGG TGCTGTTGGG AGTGTAGCAG TGAGGACGAC CAGAGGTCAC  120

TCTTGTCGCC ATCTTGGTTT TG ATG GGT TTT ACT GGC TTC TTT ACT GCA ACC   172
                         Met Gly Phe Thr Gly Phe Phe Thr Ala Thr
                                         -25                 -20

TGT TTT ATC AGC AAG GTC TTT ATG ACC TGT ATC TTG TGC AGA CCT CCT    220
Cys Phe Ile Ser Lys Val Phe Met Thr Cys Ile Leu Cys Arg Pro Pro
                -15                 -10                  -5

ATC TCA TCC TGT GTC TTA GAA TGC GGG                                247
Ile Ser Ser Cys Val Leu Glu Cys Gly
              1               5
```

(2) INFORMATION FOR SEQ ID NO: 183:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 383 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(61..347)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 88..374
id W84548
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(1..48)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 387..434
id W84548
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(61..282)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 154..375
id N66815
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(279..347)

(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 88..156
id N66815
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(12..48)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 388..424
id N66815
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(89..282)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 155..348
id N24160
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(279..347)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 89..157
id N24160
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(200..347)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 88..235
id N66833
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(239..347)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 11..119
id H88111
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(208..239)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 120..151
id H88111
est

(ix) FEATURE:

```
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: 258..362
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 4.4
                                seq LIVLLPVLFFSLK/NF

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 183:

AACACAAATA GCACTGTCAC CTCTAATATG AACATTAGTT TGAGGTAGTT TTTTTCTAAA   60

GCAAAAATTT TAACTGTTTT CTAATTGTCA AGCACTATTT TCATTAAAAG TGTCTAATGA  120

ATCATGATAT ACTCTTCCAT TTGTTGTGTC TATTTTTTAT ATATTTGGTA TTTTTTGAAA  180

ATTCCAAATA CTCATGTCTC AAGTAAGCTT AAACTACAAC TTGTCACATA AAGGAAGTCT  240

TAAGTGGAGT TCACAGA ATG ATA ATG TAT CTA TTT GTC ATT TGT GTT ATA      290
                   Met Ile Met Tyr Leu Phe Val Ile Cys Val Ile
                   -35                 -30                 -25

TTT GAA ATT ATT AGA AAT TAT GCT TTT TCC ATT TTA ATT GTA TTG CTG     338
Phe Glu Ile Ile Arg Asn Tyr Ala Phe Ser Ile Leu Ile Val Leu Leu
                -20                 -15                 -10

CCA GTG CTA TTT TTT TCT TTA AAA AAT TTT ATT CTT AGC ACA CAG         383
Pro Val Leu Phe Phe Ser Leu Lys Asn Phe Ile Leu Ser Thr Gln
             -5                   1               5


(2) INFORMATION FOR SEQ ID NO: 184:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 349 base pairs
        (B) TYPE: NUCLEIC ACID
        (C) STRANDEDNESS: DOUBLE
        (D) TOPOLOGY: LINEAR

    (ii) MOLECULE TYPE: CDNA

    (vi) ORIGINAL SOURCE:
        (A) ORGANISM: Homo Sapiens
        (F) TISSUE TYPE: Substantia nigra

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 63..281
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 96
                                region 1..219
                                id R06709
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 166..336
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 96
                                region 104..274
                                id W92876
```

est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 63..165
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 2..104
id W92876
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 165..351
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 63..249
id N31364
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 102..165
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..64
id N31364
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 166..351
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 64..249
id N75642
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 105..165
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 4..64
id N75642
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(269..324)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 259..314
id W92774
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(322..351)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93

region 231..260
id W92774
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 176..286
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.3
seq LWEKLTLLSPGIA/VT

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 184:

```
AAGTAAACAT GAGCCTCTGG TAGATAAGAG AAACCGCGAT CGGAGTACGG CGCGTGCGWG   60

NATCAGGGAT CGCGATTGCG AATCCTCCGC TGAGGTGATT TGGATATCCC TAGAACGTTG  120

AGGGCACGAG TCGGGTCCTG AGACCAGGTC CTCAGCCAGC AGAGCCACGT TCCTT ATG   178
                                                             Met

AGC ACC GTG GGT TTA TTK CAT TTT CCT ASR CCA CTG ACC CGA ATA TGC   226
Ser Thr Val Gly Leu Xaa His Phe Pro Xaa Pro Leu Thr Arg Ile Cys
    -35                 -30                 -25

CCG GCG CCA TGG GGA CTC CGG CTT TGG GAG AAG CTG ACG TTG TTA TCC   274
Pro Ala Pro Trp Gly Leu Arg Leu Trp Glu Lys Leu Thr Leu Leu Ser
-20                 -15                 -10                  -5

CCA GGA ATA GCT GTS ACT CCG GTC CAG ATG GCA GGC AAG AAG GAC TAS   322
Pro Gly Ile Ala Val Thr Pro Val Gln Met Ala Gly Lys Lys Asp Xaa
                  1               5                  10

CCT GCA CTG CTT TCC TTG GAT GAG AAT                               349
Pro Ala Leu Leu Ser Leu Asp Glu Asn
        15                  20
```

(2) INFORMATION FOR SEQ ID NO: 185:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 268 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 64..107
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 336..379
id AA100380
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 32..70
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.3
seq MLALAXHLSTVES/EK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 185:

```
CATTGAGCTC GGGCTGAGTG AGGCCCAGGT G ATG CTG GCT CTA GCC ASN CAC     52
                                   Met Leu Ala Leu Ala Xaa His
                                               -10

CTG AGC ACA GTG GAG TCG GAG AAA CAG AAG CTG CGG GCT CAG GTG CGG   100
Leu Ser Thr Val Glu Ser Glu Lys Gln Lys Leu Arg Ala Gln Val Arg
     -5                   1               5                  10

CGG CTA TGC CAG GAG AAC CAG TGG CTG CGG GAT GAG CTG GCT GGC ACC   148
Arg Leu Cys Gln Glu Asn Gln Trp Leu Arg Asp Glu Leu Ala Gly Thr
                 15                  20                  25

CAG CAG CGG CTA CAG CGC AGT GAA CAG GCT GTG GCT CAG CTG GAG GAG   196
Gln Gln Arg Leu Gln Arg Ser Glu Gln Ala Val Ala Gln Leu Glu Glu
             30                  35                  40

GAA AAG AAG CAC CTG GAG TTC CTG GGG CAG CTG CGG CAG TAT GAT GAG   244
Glu Lys Lys His Leu Glu Phe Leu Gly Gln Leu Arg Gln Tyr Asp Glu
         45                  50                  55

GAT GGA CAT ACC TCG GAG GCG GGG                                   268
Asp Gly His Thr Ser Glu Ala Gly
     60                  65
```

(2) INFORMATION FOR SEQ ID NO: 186:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 305 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 14..122
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..109
id HUM429E03B
est

```
(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 114..230
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 95
                             region 100..216
                             id HUM429E03B
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 226..307
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 96
                             region 211..292
                             id HUM429E03B
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 114..274
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 92
                             region 85..245
                             id T31768
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 29..122
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 98
                             region 1..94
                             id T31768
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 262..307
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 95
                             region 234..279
                             id T31768
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 103..307
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 93
                             region 95..299
                             id T80259
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 21..120
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 95
                             region 15..114
                             id T80259
```

est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 114..307
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 68..261
id N32697
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 45..87
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 1..43
id N32697
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 92..122
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 47..77
id N32697
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 114..307
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 59..252
id N44613
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 55..122
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 1..68
id N44613
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 147..248
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.3
seq QLFAFLNLLPVEA/DI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 186:

AGTCGTCCCT GCTAGTACTC CGGGCTGTGG GGGTCGGTGC GGATATTCAG TCATGAAATC 60

AGGGTAGGGA CTTCTCCCGC AGCGACGCGG CTGGCAAGAC TGTTTGTGTT GCGGGGGCCG 120

```
GAACTTCAAG GTGATTTTAC AACGAG ATG CTG CTC TCC ATA GGG ATG CTC ATG    173
                             Met Leu Leu Ser Ile Gly Met Leu Met
                                             -30

CTG TSA GCC ACA CAA GTC TAC ACC ATC TTG ACT GTC CAG CTC TTT GCA    221
Leu Xaa Ala Thr Gln Val Tyr Thr Ile Leu Thr Val Gln Leu Phe Ala
-25                 -20                 -15                 -10

TTC TTA AAC CTA CTG CCT GTA GAA GCA GAC ATT TTA GCA TAT AAC TTT    269
Phe Leu Asn Leu Leu Pro Val Glu Ala Asp Ile Leu Ala Tyr Asn Phe
                 -5                   1                 5

GAA AAT GCA TCT CAG ACA TTT GAT GAC CTC CCC GCA                    305
Glu Asn Ala Ser Gln Thr Phe Asp Asp Leu Pro Ala
        10                  15
```

(2) INFORMATION FOR SEQ ID NO: 187:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 333 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 95..297
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 63..265
id N31364
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 32..95
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..64
id N31364
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 293..329
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 262..298
id N31364
est

(ix) FEATURE:

(A) NAME/KEY: other
(B) LOCATION: 2..211
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 10..219
id R06709
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 203..235
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 212..244
id R06709
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 96..329
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 64..297
id N75642
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 35..95
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 4..64
id N75642
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 96..266
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 104..274
id W92876
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 2..95
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 11..104
id W92876
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(252..329)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 183..260
id W92774
est

```
     (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: complement(199..254)
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 100
                                  region 259..314
                                  id W92774
                                  est

     (ix) FEATURE:
          (A) NAME/KEY: sig_peptide
          (B) LOCATION: 106..216
          (C) IDENTIFICATION METHOD: Von Heijne matrix
          (D) OTHER INFORMATION:  score 4.3
                                  seq LWEKLTLLSPGIA/VT

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 187:

AGCGATTGCG AATCCTCCGC TGAGGTGATT TGGATATCCC TAGAACGTTG AGGGCACGAG    60

TCGGGTCCTG AGACCAGGTC CTCAGCCAGC AGAGCCACGT TCCTT ATG AGC ACC GTG   117
                                                  Met Ser Thr Val
                                                          -35

GGT TTA TTT CAT TTT CCT ACA CCA CTG ACC CGA ATA TGC CCG GCG CCA     165
Gly Leu Phe His Phe Pro Thr Pro Leu Thr Arg Ile Cys Pro Ala Pro
            -30                 -25                 -20

TGG GGA CTC CGG CTT TGG GAG AAG CTG ACG TTG TTA TCC CCA GGA ATA     213
Trp Gly Leu Arg Leu Trp Glu Lys Leu Thr Leu Leu Ser Pro Gly Ile
        -15                 -10                  -5

GCT GTC ACT CCG GTC CAG ATG GCA GGC AAG AAG GAC TAC CCT GCA CTG     261
Ala Val Thr Pro Val Gln Met Ala Gly Lys Lys Asp Tyr Pro Ala Leu
      1               5                  10                  15

CTT TCC TTG GAT GAG VAT GAA CTC GAA GAG CAG TTT GTG AAA GGA CAC     309
Leu Ser Leu Asp Glu Xaa Glu Leu Glu Glu Gln Phe Val Lys Gly His
                 20                  25                  30

GGT CCA GGG GGC CAG GCA ACG CGG                                     333
Gly Pro Gly Gly Gln Ala Thr Arg
             35


(2) INFORMATION FOR SEQ ID NO: 188:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 510 base pairs
          (B) TYPE: NUCLEIC ACID
          (C) STRANDEDNESS: DOUBLE
          (D) TOPOLOGY: LINEAR

     (ii) MOLECULE TYPE: CDNA

     (vi) ORIGINAL SOURCE:
          (A) ORGANISM: Homo Sapiens
          (F) TISSUE TYPE: Substantia nigra
```

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 188..485
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 171..468
id R76663
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 17..186
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..170
id R76663
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 183..367
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 169..353
id H67124
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 14..185
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..172
id H67124
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 214..430
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 81..297
id R53683
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 133..227
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 1..95
id R53683
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 16..114
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.2
seq LLNFLGLWSWICK/KW

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 188:

```
ATTCCCTTCT GAGCA ATG GAG CTT ACC ATC TTT ATC CTG AGA CTG GCC ATT    51
                 Met Glu Leu Thr Ile Phe Ile Leu Arg Leu Ala Ile
                             -30                 -25

TAC ATC CTG ACA TTT CCC TTG TAC CTG CTG AAC TTT CTG GGC TTG TGG     99
Tyr Ile Leu Thr Phe Pro Leu Tyr Leu Leu Asn Phe Leu Gly Leu Trp
    -20                 -15                 -10

AGC TGG ATA TGC AAA AAA TGG TTC CCC TAC TTC TTG GTG AGG TTC ACT    147
Ser Trp Ile Cys Lys Lys Trp Phe Pro Tyr Phe Leu Val Arg Phe Thr
 -5                   1               5                  10

GTG ATA TAC AAC GAA CAG ATG GCA AGC AAG AAG CGG GAG CTC TTC AGT    195
Val Ile Tyr Asn Glu Gln Met Ala Ser Lys Lys Arg Glu Leu Phe Ser
             15                  20                  25

AAC CTG CAG GAG TTT GCG GGC CCC TCC GGG AAA CTC TCC CTG CTG GAA    243
Asn Leu Gln Glu Phe Ala Gly Pro Ser Gly Lys Leu Ser Leu Leu Glu
        30                  35                  40

GTG GGC TGT GGC ACG GGG GCC AAC TTC AAG TTC TAC CCA CCT GGG TGC    291
Val Gly Cys Gly Thr Gly Ala Asn Phe Lys Phe Tyr Pro Pro Gly Cys
    45                  50                  55

AGG GTG ACC TGT ATT GAC CCC AAC CCC AAC TTT GAG AAG TTT TTG ATC    339
Arg Val Thr Cys Ile Asp Pro Asn Pro Asn Phe Glu Lys Phe Leu Ile
 60                  65                  70                  75

AAG AGC ATT GCA GAG AAC CGA CAC CTG CAG TTT GAG CGC TTT GTG GTA    387
Lys Ser Ile Ala Glu Asn Arg His Leu Gln Phe Glu Arg Phe Val Val
                 80                  85                  90

GCT GCC GGG GAG AAC ATG CAC CAG GTG GCT GAT GGC TCT GTG GAT GTG    435
Ala Ala Gly Glu Asn Met His Gln Val Ala Asp Gly Ser Val Asp Val
             95                 100                 105

GTG GTC TGC ACC CTG GTG CTG TGC TCT GTG AAG AAC CAG GAG CGG ATT    483
Val Val Cys Thr Leu Val Leu Cys Ser Val Lys Asn Gln Glu Arg Ile
        110                 115                 120

CTC CGC GAG GTG TGC AGA GTG CTG AGA                                510
Leu Arg Glu Val Cys Arg Val Leu Arg
    125                 130
```

(2) INFORMATION FOR SEQ ID NO: 189:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 155 base pairs
  (B) TYPE: NUCLEIC ACID
  (C) STRANDEDNESS: DOUBLE
  (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(2..122)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 9..129
id R16550
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 85..150
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..66
id N42850
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 9..119
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.1
seq LLLYLCCMINIHH/LP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 189:

```
ATTGAACT ATG AGT CTC CTG CAT GGC AAC AAA ATG TGT GTC ACC ATC AGG      50
         Met Ser Leu Leu His Gly Asn Lys Met Cys Val Thr Ile Arg
                 -35                 -30                 -25

CCA ACA GGC CAG CCC TTG AAT GGG GAT TTA TTA CTG TTG TAT CTA TGT       98
Pro Thr Gly Gln Pro Leu Asn Gly Asp Leu Leu Leu Leu Tyr Leu Cys
            -20                 -15                 -10

TGC ATG ATA AAC ATT CAT CAC CTT CCT CCT GTA GTC CTG CCT CGT ACT      146
Cys Met Ile Asn Ile His His Leu Pro Pro Val Val Leu Pro Arg Thr
         -5                  1               5

CCC CAA GGG                                                          155
Pro Gln Gly
 10
```

(2) INFORMATION FOR SEQ ID NO: 190:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 250 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens

(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(90..123)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 1..34
id H04995
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 122..172
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.9
seq ISYFIAFPNLSQA/EL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 190:

```
ATAATCATAT TTCAAAATGA ATAGCAAAGA GGTTTATAAT CAAGTTTTAT AAAAATTCCA    60

AATGTAATAA AGTTATATTT GTAACTTACA TATACTGCAA AAATGGTAGT GATTCAAATG   120

T ATG TCT TTC AAT ATA TCG TAT TTT ATT GCC TTT CCA AAT CTC TCC CAG   169
  Met Ser Phe Asn Ile Ser Tyr Phe Ile Ala Phe Pro Asn Leu Ser Gln
      -15                 -10                     -5

GCA GAA TTA ACA CAT CCC AGG TGC TCT TAC ACA GGG TTG AGT AGT TCC     217
Ala Glu Leu Thr His Pro Arg Cys Ser Tyr Thr Gly Leu Ser Ser Ser
    1               5                  10                  15

TGT GGA TTT CAG TTG AGT GAT ACC CCC CAC AGG                         250
Cys Gly Phe Gln Leu Ser Asp Thr Pro His Arg
                20                  25
```

(2) INFORMATION FOR SEQ ID NO: 191:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 379 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 193..356
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 131..294

id AA148880
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 93..193
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 30..130
id AA148880
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 63..97
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 91
region 1..35
id AA148880
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 122..307
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 131..316
id H92506
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 9..97
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 1..89
id H92506
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 302..361
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq LTIILLPVHLLIT/IY

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 191:

```
ACTTTTTGCG AGCTTTCCGA GTGCCAGGCG CCGGCCGGCT GCGAAGACGC GGTGGGCCGC    60

CCCTCCGATT GAAATCACAG AAGATATTCG KGKTCCTTCT TAAGAGAAAA AGAGGACATT   120

TGCGTACTTT ATTGTCGGCT TCCAAAGATT ACTAACTTTT ATCTGTATCA CTAAGATTGA   180

ACTGCCTTGG CTGTACTGCT ATTCTTACTG CTGCTTCTAT TATTGCCTTC TTCAGCACAA   240

TAAGGCTTTC AAAAGCCAAA GAATAACAAG AAATAAGCAC CATTTTAGAA GCCTTTCCAC   300

T ATG AAA CTT AAG CWA AAT GTG CTC ACC ATT ATT TTG CTG CCT GTC CAC   349
  Met Lys Leu Lys Xaa Asn Val Leu Thr Ile Ile Leu Leu Pro Val His
```

```
 -20                  -15                   -10                   -5

TTG TTA ATA ACA ATA TAC AGT GCC CTT ATA                           379
Leu Leu Ile Thr Ile Tyr Ser Ala Leu Ile
                 1               5
```

(2) INFORMATION FOR SEQ ID NO: 192:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 176 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 25..173
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 91
region 67..215
id H46464
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 73..173
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 44..144
id C17500
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 32..67
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 1..36
id C17500
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(112..173)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 215..276
id AA143237
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(65..117)

(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 270..322
id AA143237
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(125..173)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 241..289
id W57931
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(8..50)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 368..410
id W57931
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 9..50
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq AALVTVLFTGVRR/LH

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 192:

```
AGCTCAGC ATG GCT GCT TTA GTG ACT GTT CTC TTC ACA GGT GTC CGG AGG      50
         Met Ala Ala Leu Val Thr Val Leu Phe Thr Gly Val Arg Arg
                         -10                 -5

CTG CAC TGC AGC GCR SCG CTT GGG CGG GCG GCC AGT GGC GRC TAC AGC       98
Leu His Cys Ser Ala Xaa Leu Gly Arg Ala Ala Ser Gly Xaa Tyr Ser
  1             5                  10                  15

AGG AAC TGG CTG CCA ACC CCT CCG GCT ACG GGC CCC TTA CCG AGC TCC      146
Arg Asn Trp Leu Pro Thr Pro Pro Ala Thr Gly Pro Leu Pro Ser Ser
            20                  25                  30

CAG ACT GGT CAT ATG CGG ATG GCC GCC AGG                              176
Gln Thr Gly His Met Arg Met Ala Ala Arg
        35                  40
```

(2) INFORMATION FOR SEQ ID NO: 193:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 236 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Surrenals

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 85..236
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 91..242
id AA031580
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 3..87
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 10..94
id AA031580
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 85..236
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 82..233
id W80981
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 3..87
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..85
id W80981
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 90..236
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 89..235
id R69999
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 3..88
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 3..88
id R69999
est

(ix) FEATURE:
(A) NAME/KEY: other

```
(B) LOCATION: 90..236
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION:  identity 97
                        region 89..235
                        id R76832
                        est

(ix) FEATURE:
   (A) NAME/KEY: other
   (B) LOCATION: 3..88
   (C) IDENTIFICATION METHOD: blastn
   (D) OTHER INFORMATION:  identity 100
                           region 3..88
                           id R76832
                           est

(ix) FEATURE:
   (A) NAME/KEY: other
   (B) LOCATION: 90..236
   (C) IDENTIFICATION METHOD: blastn
   (D) OTHER INFORMATION:  identity 97
                           region 89..235
                           id R80120
                           est

(ix) FEATURE:
   (A) NAME/KEY: other
   (B) LOCATION: 3..88
   (C) IDENTIFICATION METHOD: blastn
   (D) OTHER INFORMATION:  identity 100
                           region 3..88
                           id R80120
                           est

(ix) FEATURE:
   (A) NAME/KEY: sig_peptide
   (B) LOCATION: 6..131
   (C) IDENTIFICATION METHOD: Von Heijne matrix
   (D) OTHER INFORMATION:  score 3.7
                           seq ILMRDFSPSGIFG/AF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 193:

ATAAA ATG GCG TCA GTT GGT GAG TGT CCG GCC CCA GTA CCA GTG AAG GAC    50
      Met Ala Ser Val Gly Glu Cys Pro Ala Pro Val Pro Val Lys Asp
              -40                 -35                 -30

AAG AAA CTT CTG GAG GTC AAA CTG GGG GAG CTG CCA AGC TGG ATC TTG      98
Lys Lys Leu Leu Glu Val Lys Leu Gly Glu Leu Pro Ser Trp Ile Leu
        -25                 -20                 -15

ATG CGG GAC TTC AGT CCT AGT GGC ATT TTC GGA GCG TTT CAA AGA GGT     146
Met Arg Asp Phe Ser Pro Ser Gly Ile Phe Gly Ala Phe Gln Arg Gly
    -10                 -5                   1               5

TAC TAS CGG TAC TAC AAC VAG TAC ATC AAT GTG RAG AAG GGG AGC ATC     194
Tyr Xaa Arg Tyr Tyr Asn Xaa Tyr Ile Asn Val Xaa Lys Gly Ser Ile
                10                  15                  20

TCG GGG ATT AMC ATG GTG CTG GCA TGC TAC GTG CTC TTT AGC             236
```

```
Ser Gly Ile Xaa Met Val Leu Ala Cys Tyr Val Leu Phe Ser
            25                  30                  35
```

(2) INFORMATION FOR SEQ ID NO: 194:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 155 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Surrenals

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 61..155
(C) IDENTIFICATION METHOD: fasta
(D) OTHER INFORMATION: identity 98
region 4..98
id HUMPASP
vrt

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 77..155
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..79
id W44779
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 81..125
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 10.9
seq ALLVLVTVALASA/HH

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 194:

```
AGCAATCAGG AGAAGAAAGC AAAGAACACT CAGGATTATA AAAGCAGATG AGACCTACCC   60

ACTAGACCTG GTCAGACACA ATG TTG GCA CTC TTG GTT CTG GTG ACT GTG GCC  113
                      Met Leu Ala Leu Leu Val Leu Val Thr Val Ala
                      -15                 -10                  -5

CTG GCA TCT GCT CAT CAT GGT GGT GAG CAC TTT GAA GGC GCG            155
Leu Ala Ser Ala His His Gly Gly Glu His Phe Glu Gly Ala
                1               5                  10
```

(2) INFORMATION FOR SEQ ID NO: 195:

```
(i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 214 base pairs
      (B) TYPE: NUCLEIC ACID
      (C) STRANDEDNESS: DOUBLE
      (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
      (A) ORGANISM: Homo Sapiens
      (F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 47..211
      (C) IDENTIFICATION METHOD: fasta
      (D) OTHER INFORMATION:  identity 98
                              region 1..165
                              id HSU16129
                              vrt

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 3..211
      (C) IDENTIFICATION METHOD: fasta
      (D) OTHER INFORMATION:  identity 92
                              region 117..326
                              id RATGLUR4A
                              vrt

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 35..211
      (C) IDENTIFICATION METHOD: fasta
      (D) OTHER INFORMATION:  identity 91
                              region 1..177
                              id GGGLUR4
                              vrt

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 47..211
      (C) IDENTIFICATION METHOD: fasta
      (D) OTHER INFORMATION:  identity 93
                              region 1..165
                              id RATAMPASGD
                              vrt

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 47..211
      (C) IDENTIFICATION METHOD: fasta
      (D) OTHER INFORMATION:  identity 93
                              region 1..165
                              id RATGLURD
                              vrt

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 47..211
      (C) IDENTIFICATION METHOD: fasta
```

```
        (D) OTHER INFORMATION:  identity 93
                                region 1..165
                                id S94371
                                vrt

   (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 61..211
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 100
                                region 93..243
                                id R93859
                                est

   (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 136..210
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 98
                                region 1..75
                                id R16185
                                est

   (ix) FEATURE:
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: 47..109
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 8.9
                                seq VLLFSGFWGLAMG/AF

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 195:

AAGAGAGAAA GAGAGAGAGC GCGCGCCAGG GAGAGGAGAA AAGAAG ATG AGG ATT       55
                                                   Met Arg Ile
                                                       -20

ATT TCC AGA CAG ATT GTC TTG TTA TTT TCT GGA TTT TGG GGA CTC GCC     103
Ile Ser Arg Gln Ile Val Leu Leu Phe Ser Gly Phe Trp Gly Leu Ala
            -15                 -10                 -5

ATG GGA GCC TTT CCG AGC AGC GTG CAA ATA GGT GGT CTC TTC ATC CGA     151
Met Gly Ala Phe Pro Ser Ser Val Gln Ile Gly Gly Leu Phe Ile Arg
         1               5                  10

AAC ACA GAT CAG GAA TAC ACT GCT TTT CGA TTA GCA ATT TTT CTT CAT     199
Asn Thr Asp Gln Glu Tyr Thr Ala Phe Arg Leu Ala Ile Phe Leu His
 15                  20                  25                  30

AAC ACC AGC CCC GGG                                                 214
Asn Thr Ser Pro Gly
                 35


(2) INFORMATION FOR SEQ ID NO: 196:

   (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 29 amino acids
        (B) TYPE: AMINO ACID
        (D) TOPOLOGY: LINEAR
```

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 14.3
seq LLLCAVLLSLASA/SS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 196:

```
Met Arg Val Arg Ile Gly Leu Thr Leu Leu Leu Cys Ala Val Leu Leu
    -20                 -15                 -10

Ser Leu Ala Ser Ala Ser Ser Asp Glu Glu Gly Asn Gly
 -5                 1               5
```

(2) INFORMATION FOR SEQ ID NO: 197:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -28..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 11.1
seq GLLFLLLLLMLLA/DP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 197:

```
Met Gly Leu His Leu Arg Pro Tyr Arg Val Gly Leu Leu Pro Asp Gly
            -25                 -20                 -15

Leu Leu Phe Leu Leu Leu Leu Leu Met Leu Leu Ala Asp Pro Ala Leu
        -10                 -5                  1

Pro Ala Ala Arg
  5
```

(2) INFORMATION FOR SEQ ID NO: 198:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 134 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 10.6
seq LLLGAVSWPPASA/SG

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 198:

```
Met Ala Thr Leu Ser Phe Val Phe Leu Leu Leu Gly Ala Val Ser Trp
    -20                 -15                 -10

Pro Pro Ala Ser Ala Ser Gly Gln Glu Phe Trp Pro Gly Gln Ser Ala
 -5                  1               5                  10

Ala Asp Ile Leu Ser Gly Ala Ala Ser Arg Arg Arg Tyr Leu Leu Tyr
             15                  20                  25

Asp Val Asn Pro Pro Glu Gly Phe Asn Leu Arg Arg Asp Val Tyr Ile
        30                  35                  40

Arg Ile Ala Ser Leu Leu Lys Thr Leu Leu Lys Thr Glu Glu Trp Val
    45                  50                  55

Leu Xaa Leu Pro Pro Trp Gly Arg Leu Xaa Xaa Trp Gln Ser Xaa Asp
60                  65                  70                  75

Ile His Gln Val Arg Ile Pro Trp Ser Glu Phe Phe Asp Leu Pro Ser
                80                  85                  90

Leu Asn Lys Asn Ile Pro Val Ile Glu Tyr Glu Gln Phe Ile Ala Glu
            95                  100                 105

Ser Gly Gly Pro Phe Ile
        110
```

(2) INFORMATION FOR SEQ ID NO: 199:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 59 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -54..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9.1
seq VLCLRGLVSLAFQ/GP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 199:

```
Met Phe Leu Phe Leu Ser Pro Ala Thr Pro Val Leu Pro Pro Ser Leu
                -50                 -45                 -40

Asp Ser Arg Asp Leu Leu Pro His Leu Phe Trp Gly Arg Ala Gly Ser
            -35                 -30                 -25

Ser Ser Ser Ser Pro Ala Leu Ser Pro Val Leu Cys Leu Arg Gly Leu
        -20                 -15                 -10

Val Ser Leu Ala Phe Gln Gly Pro His Pro Glu
     -5                  1               5
```

(2) INFORMATION FOR SEQ ID NO: 200:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 70 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.8
seq LLWALLFMQSLWP/QL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 200:

```
Met Trp Ala Met Glu Ser Gly His Leu Leu Trp Ala Leu Leu Phe Met
    -20                 -15                 -10

Gln Ser Leu Trp Pro Gln Leu Thr Asp Gly Ala Thr Arg Val Tyr Tyr
 -5                  1               5                  10

Leu Gly Ile Arg Asp Val Gln Trp Asn Tyr Ala Pro Lys Gly Arg Asn
            15                  20                  25
```

Val Ile Thr Asn Gln Pro Leu Asp Ser Asp Ile Val Ala Ser Ser Phe
30 35 40

Leu Lys Ser Asp Asp Gly
45

(2) INFORMATION FOR SEQ ID NO: 201:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -19..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.7
seq ILGLLCCVLATMA/NP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 201:

Met Thr His Tyr Arg Asn Ile Leu Gly Leu Leu Cys Cys Val Leu Ala
-15 -10 -5

Thr Met Ala Asn Pro Gly
1

(2) INFORMATION FOR SEQ ID NO: 202:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -15..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.4
seq LLLLLASLIERSS/KT

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 202:

```
Met Lys Leu Leu Leu Leu Leu Ala Ser Leu Ile Glu Arg Ser Ser Lys
-15                 -10                 -5                    1

Thr Ser Cys Xaa Xaa Gln His Tyr Ser Ser Gln
             5                  10
```

(2) INFORMATION FOR SEQ ID NO: 203:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 52 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -38..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.2
seq SFXLFLALCASFS/FF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 203:

```
Met Ala Arg Asn Gln Ala Leu Val Cys Leu Pro Ser Phe Gln Asn Ala
            -35                 -30                 -25

Phe Ile Pro Val Glu Asp Leu Pro Thr Ser Phe Xaa Leu Phe Leu Ala
        -20                 -15                 -10

Leu Cys Ala Ser Phe Ser Phe Phe Leu Xaa Leu Ser Leu Ser Leu Pro
     -5                   1               5                  10

Ser Phe Phe Phe
```

(2) INFORMATION FOR SEQ ID NO: 204:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide

(B) LOCATION: -35..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.2
seq SLLLLFYSFYVLA/VK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 204:

```
Met Pro Asn Glu Ser Trp Gln Ile Pro Cys Gly Lys Gln Glu Ala Glu
-35                 -30                 -25                 -20

Thr Leu Phe Asn Phe Gln Ser Leu Leu Leu Leu Phe Tyr Ser Phe Tyr
                -15                 -10                 -5

Val Leu Ala Val Lys Arg Gly Pro Gly
            1               5
```

(2) INFORMATION FOR SEQ ID NO: 205:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -36..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8
seq LVLLICLVSSYLP/QL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 205:

```
Met Gln Thr Thr Phe Ile Asp Val Thr Val Asp Gln His Val Ala Lys
    -35                 -30                 -25

Ser Asn Asp His Leu Ser Val Leu Val Leu Leu Ile Cys Leu Val Ser
-20                 -15                 -10                 -5

Ser Tyr Leu Pro Gln Leu Pro
                1
```

(2) INFORMATION FOR SEQ ID NO: 206:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 80 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -73..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.8
seq YLPLLAGLGLTLA/AP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 206:

```
Met Ala Leu Gly Glu Glu Lys Ala Glu Ala Glu Ala Ser Xaa Asp Thr
            -70                 -65                 -60

Lys Ala Gln Ser Tyr Gly Arg Gly Ser Cys Arg Glu Arg Glu Leu Asp
        -55                 -50                 -45

Ile Pro Gly Pro Met Ser Gly Glu Gln Pro Pro Arg Leu Glu Ala Glu
    -40                 -35                 -30

Gly Gly Leu Ile Ser Pro Val Trp Gly Ala Glu Xaa Tyr Leu Pro Leu
-25                 -20                 -15                 -10

Leu Ala Gly Leu Gly Leu Thr Leu Ala Ala Pro Leu Glu Pro Thr Thr
                -5                  1               5
```

(2) INFORMATION FOR SEQ ID NO: 207:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.7
seq LLCISPFVPFTSG/NK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 207:

```
Met Thr Ser Leu Tyr Leu Lys His Leu Leu Cys Ile Ser Pro Phe Val
    -20                 -15                 -10

Pro Phe Thr Ser Gly Asn Lys Leu Tyr Tyr Thr Met Ile Tyr Trp Leu
 -5                  1               5                   10
```

```
Phe Lys Thr Val Leu Asn Met His Gly
            15                  20


(2) INFORMATION FOR SEQ ID NO: 208:

      (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 58 amino acids
            (B) TYPE: AMINO ACID
            (D) TOPOLOGY: LINEAR

      (ii) MOLECULE TYPE: PROTEIN

      (vi) ORIGINAL SOURCE:
            (A) ORGANISM: Homo Sapiens
            (F) TISSUE TYPE: Cerebellum

      (ix) FEATURE:
            (A) NAME/KEY: sig_peptide
            (B) LOCATION: -47..-1
            (C) IDENTIFICATION METHOD: Von Heijne matrix
            (D) OTHER INFORMATION:  score 7.6
                                    seq CLATLTLFHTSFS/FQ

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 208:

Met Thr Asp Ser Pro Asn Ala His Gly Leu Ala Leu Thr Thr Lys Trp
        -45                 -40                 -35

Met Met Pro Ala Val Ser Leu Asn Leu Thr Tyr Tyr Leu Pro Ser Trp
    -30                 -25                 -20

Tyr Leu Cys Leu Ala Thr Leu Thr Leu Phe His Thr Ser Phe Ser Phe
-15                 -10                 -5                   1

Gln Ala Ser Glu Ser Val Lys Ala Ile Thr
            5                   10


(2) INFORMATION FOR SEQ ID NO: 209:

      (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 78 amino acids
            (B) TYPE: AMINO ACID
            (D) TOPOLOGY: LINEAR

      (ii) MOLECULE TYPE: PROTEIN

      (vi) ORIGINAL SOURCE:
            (A) ORGANISM: Homo Sapiens
            (F) TISSUE TYPE: Substantia nigra

      (ix) FEATURE:
            (A) NAME/KEY: sig_peptide
            (B) LOCATION: -48..-1
            (C) IDENTIFICATION METHOD: Von Heijne matrix
            (D) OTHER INFORMATION:  score 7.5
```

seq FVILLLFIFTVVS/LV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 209:

```
Met Ala Ser Ser His Trp Asn Glu Thr Thr Thr Ser Val Tyr Gln Tyr
            -45                 -40                 -35

Leu Gly Phe Gln Val Gln Lys Ile Tyr Pro Phe His Asp Asn Trp Asn
        -30                 -25                 -20

Thr Ala Cys Phe Val Ile Leu Leu Leu Phe Ile Phe Thr Val Val Ser
    -15                 -10                  -5

Leu Val Val Leu Ala Phe Leu Tyr Glu Val Leu Xaa Xaa Cys Cys Cys
  1               5                  10                  15

Val Lys Asn Lys Thr Val Lys Asp Leu Lys Ser Glu Pro Lys
             20                  25                  30
```

(2) INFORMATION FOR SEQ ID NO: 210:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -25..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.4
seq IFLLNMWVACLLS/GE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 210:

```
Met Ser Leu Leu Phe Val Phe Cys Leu Glu Cys Ser Ile Phe Leu Leu
-25                 -20                 -15                 -10

Asn Met Trp Val Ala Cys Leu Leu Ser Gly Glu Ile Pro His Ser Ser
                 -5                   1                   5

Trp Xaa Xaa Lys Leu Ile Gly Thr Leu Pro Thr Ser
         10                  15
```

(2) INFORMATION FOR SEQ ID NO: 211:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids

(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -18..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.4
seq VLLLLPLVAFITL/KF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 211:

```
Met Phe Val Val Thr Val Leu Leu Leu Leu Pro Leu Val Ala Phe Ile
            -15                 -10                 -5

Thr Leu Lys Phe Cys Asn Leu Ile Asn Phe Pro Thr Xaa Arg His Gly
         1               5                   10
```

(2) INFORMATION FOR SEQ ID NO: 212:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 67 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -22..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.2
seq IIYALQFLFLVFA/PS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 212:

```
Met Asn Arg Ser Cys Arg Asn Thr Gly Ile Ile Tyr Ala Leu Gln Phe
        -20                 -15                 -10

Leu Phe Leu Val Phe Ala Pro Ser Ser Leu Gly Tyr Phe Glu Trp Ile
    -5                   1               5                  10

Val Ala Ile Asn Gln Asp Leu Val Leu Phe Val Phe Cys Leu Ser Phe
                15                  20                  25

Ser Leu Arg Ile Ser Ile Ile Gln Gly Lys Arg Lys Ala Ala Phe Pro
            30                  35                  40
```

```
Thr Pro Pro
        45
```

(2) INFORMATION FOR SEQ ID NO: 213:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -35..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.2
seq LSLLLAWVTLTHL/LS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 213:

```
Met Thr Gln Thr Thr Trp Gly Ala Pro Thr Arg Ala Ser Asn His Pro
-35                 -30                 -25                 -20

Leu Pro Ala Trp Leu Thr Leu Ser Leu Leu Leu Ala Trp Val Thr Leu
                -15                 -10                 -5

Thr His Leu Leu Ser Val Leu Thr His Pro Thr Leu Leu
            1               5                   10
```

(2) INFORMATION FOR SEQ ID NO: 214:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 85 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -16..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.1
seq XXAVLCVCAAAWC/SQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 214:

```
Met Leu Lys Xaa Xaa Ala Val Leu Cys Val Cys Ala Ala Ala Trp Cys
    -15                 -10                 -5

Ser Gln Ser Leu Ala Ala Ala Ala Ala Val Ala Ala Xaa Gly Gly Arg
  1                 5                  10                  15

Ser Asp Gly Gly Asn Phe Leu Asp Asp Lys Gln Trp Leu Thr Thr Ile
             20                  25                  30

Ser Gln Tyr Asp Lys Glu Val Gly Gln Trp Asn Lys Phe Arg Asp Asp
         35                  40                  45

Asp Tyr Phe Arg Thr Trp Ser Pro Gly Lys Pro Phe Asp Gln Ala Leu
     50                  55                  60

Asp Xaa Ala Asn Gly
 65
```

(2) INFORMATION FOR SEQ ID NO: 215:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -20..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.9
seq LHLLGSSISPASA/SL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 215:

```
Met Ile Ser Ala His Cys Asn Leu His Leu Leu Gly Ser Ser Ile Ser
-20                 -15                 -10                  -5

Pro Ala Ser Ala Ser Leu
                  1
```

(2) INFORMATION FOR SEQ ID NO: 216:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 55 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -28..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.9
seq FLPFLLSLPLDQT/LP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 216:

```
Met Xaa Xaa Lys Ala Cys Arg Thr Leu Ala Trp Leu Pro Xaa Pro Phe
            -25                 -20                 -15

Leu Pro Phe Leu Leu Ser Leu Pro Leu Asp Gln Thr Leu Pro Arg Gln
        -10                 -5                   1

Gly Pro Gly Gln Ser Leu Ser Phe Pro Glu Asn Tyr Gln Thr Leu Pro
  5                  10                  15                  20

Lys Ser Thr Arg His Pro Gly
                 25
```

(2) INFORMATION FOR SEQ ID NO: 217:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 58 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -19..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.8
seq GLLLVFLPHPQRG/GQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 217:

```
Met Ala Val Lys Arg Leu Gly Leu Leu Leu Val Phe Leu Pro His Pro
                -15                 -10                 -5

Gln Arg Gly Gly Gln Glu Arg Ser Ala His Thr Pro Arg Gln His Pro
             1                   5                  10

Ala Arg Pro Thr Ser Leu Ser Gln Gly Glu Arg Pro Gly Arg Gly Gly
    15                  20                  25
```

```
Gly Trp Gly Asn Gly Arg Asp Ala His Gln
 30                 35
```

(2) INFORMATION FOR SEQ ID NO: 218:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 119 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -83..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.8
seq LLMILTFPFKILS/DA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 218:

```
Met Gln Ala Val Asp Asn Leu Thr Ser Ala Pro Gly Asn Thr Ser Leu
            -80                 -75                 -70

Cys Thr Arg Asp Tyr Lys Ile Thr Gln Val Leu Phe Pro Leu Leu Tyr
        -65                 -60                 -55

Thr Val Leu Phe Phe Val Gly Leu Ile Thr Asn Gly Leu Ala Met Arg
    -50                 -45                 -40

Ile Phe Phe Gln Ile Arg Ser Lys Ser Asn Phe Ile Ile Phe Leu Lys
-35                 -30                 -25                 -20

Asn Thr Val Ile Ser Asp Leu Leu Met Ile Leu Thr Phe Pro Phe Lys
                -15                 -10                  -5

Ile Leu Ser Asp Ala Lys Leu Gly Thr Gly Pro Leu Arg Thr Phe Val
              1               5                  10

Cys Gln Val Thr Ser Val Ile Phe Tyr Xaa Xaa Met Tyr Ile Ser Ile
     15                  20                  25

Ser Phe Leu Gly Leu Ile Thr
 30                 35
```

(2) INFORMATION FOR SEQ ID NO: 219:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 55 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -36..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.8
seq LWFFLPSLXCPEC/CP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 219:

```
Met Trp Thr Leu Pro Ser Leu Ser Ala Ser Phe Gln Pro Phe Leu Gly
    -35                 -30                 -25

Ser Leu Arg Pro Ser His Ile Leu Trp Phe Phe Leu Pro Ser Leu Xaa
-20                 -15                 -10                 -5

Cys Pro Glu Cys Cys Pro Pro Asp Pro Gly Ser Pro Ala Ser Arg Asp
                1               5                   10

Pro Asn Val Ala Cys Glu Arg
        15
```

(2) INFORMATION FOR SEQ ID NO: 220:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 84 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -22..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.7
seq LLLFQPSSHSATG/SS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 220:

```
Met Ser Leu Thr Asp Val Pro Met Ser Leu Leu Leu Phe Gln Pro Ser
    -20                 -15                 -10

Ser His Ser Ala Thr Gly Ser Ser Ile Lys Ile Ile Ile Leu Asn Tyr
    -5                  1               5                   10
```

```
Ile Ile Leu Gln Phe Lys Thr Leu Gln Thr Leu Pro Asn Ala Leu Arg
                15                  20                  25

Ile His Ile Lys Val Phe His Ile Tyr Cys Ser Phe Val Ser Arg Phe
            30                  35                  40

His Tyr Tyr Lys Asn Thr Ala Thr Val Phe Phe Arg Ser Val Leu Lys
        45                  50                  55

Arg Arg Met Gly
    60
```

(2) INFORMATION FOR SEQ ID NO: 221:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 72 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -17..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.6
seq LAFLLVSLYWSHM/HP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 221:

```
Met Asp Trp Ser Leu Ala Phe Leu Leu Val Ser Leu Tyr Trp Ser His
        -15                 -10                 -5

Met His Pro Cys Tyr Trp Ser Trp Pro Cys Ser Cys Gly Phe Val Asp
    1               5                   10                  15

Ser Pro Cys Ile Cys Thr Ala Ser Thr Arg Cys Cys Cys Ser Ser Leu
                20                  25                  30

Cys Cys Leu Trp Arg Leu Ala Leu Trp Asp Trp Thr Ser Asn Gly Ser
            35                  40                  45

Arg Ser Gly Ile Ala Cys Val Cys
        50                  55
```

(2) INFORMATION FOR SEQ ID NO: 222:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -15..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.5
seq LLILFFMVGRIIP/SP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 222:

```
Met Tyr Leu Leu Ile Leu Phe Phe Met Val Gly Arg Ile Ile Pro Ser
-15                 -10                  -5                   1

Pro His Arg
```

(2) INFORMATION FOR SEQ ID NO: 223:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 56 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -48..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.5
seq LLVVSCCLLFHQA/IH

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 223:

```
Met Asn Lys Pro Pro Trp Glu Glu Ser Trp Gly Gln Asn Gln Leu Ser
            -45                 -40                 -35

Gly Glu Pro Ala Thr Trp Ser Leu Cys Ile Ser Pro Leu Pro Gly Arg
        -30                 -25                 -20

Glu Pro Ser Leu Leu Val Val Ser Cys Cys Leu Leu Phe His Gln Ala
    -15                 -10                  -5

Ile His Asn Lys Leu Leu Trp Arg
  1               5
```

(2) INFORMATION FOR SEQ ID NO: 224:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -24..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.5
seq FLILLSIDSLVSG/FL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 224:

```
Met Leu Ile Leu Glu Leu Thr Met Met Leu Ser Phe Leu Ile Leu Leu
                -20                 -15                 -10

Ser Ile Asp Ser Leu Val Ser Gly Phe Leu Ser Lys Arg Lys Gly Leu
             -5                   1               5

Arg Val Cys Asp Gly Ser Arg Ser Gly
     10                  15
```

(2) INFORMATION FOR SEQ ID NO: 225:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 56 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -52..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.4
seq CLLGAAWASRLRT/QP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 225:

```
Met Lys Leu Gln Arg Ser Arg Ala Phe Arg Ile Glu Cys Ser Ala Ile
    -50                 -45                 -40
```

```
Leu Arg Arg Ala Glu Arg Leu Val Trp Asn Asp Val Cys Ser Glu Ser
    -35                 -30                 -25

Gln Ser Gln Ser Arg Asp Ser Cys Leu Leu Gly Ala Ala Trp Ala Ser
-20                 -15                 -10                 -5

Arg Leu Arg Thr Gln Pro His Pro
                1
```

(2) INFORMATION FOR SEQ ID NO: 226:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -16..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.7
seq FTLCVFTLPFLCA/CL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 226:

```
Met Val Ile Phe Thr Leu Cys Val Phe Thr Leu Pro Phe Leu Cys Ala
    -15                 -10                     -5

Cys Leu Pro Arg
  1
```

(2) INFORMATION FOR SEQ ID NO: 227:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -19..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.7
seq VLVVGTWSSQGQA/NS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 227:

```
Met Trp Gly Ala Leu Pro Val Leu Val Val Gly Thr Trp Ser Ser Gln
                -15                 -10                 -5

Gly Gln Ala Asn Ser Cys Ala Gly Arg Gly Met Gly Pro Asp Val Cys
            1               5                   10

Gly Ala
    15
```

(2) INFORMATION FOR SEQ ID NO: 228:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -16..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.7
seq LVCGFLQISLSLA/SL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 228:

```
Met Thr Arg Leu Val Cys Gly Phe Leu Gln Ile Ser Leu Ser Leu Ala
    -15                 -10                 -5

Ser Leu Phe Leu Thr Ile Pro Leu Met Trp Tyr Met Gln Ser Lys Trp
 1              5                   10                  15

Trp Arg Gly
```

(2) INFORMATION FOR SEQ ID NO: 229:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal

(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -15..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.6
seq FLLPLLLHHLTFH/GR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 229:

```
Met Asn Phe Leu Leu Pro Leu Leu Leu His His Leu Thr Phe His Gly
-15                 -10                 -5                    1

Arg Pro Leu Lys
              5
```

(2) INFORMATION FOR SEQ ID NO: 230:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 81 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -58..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.5
seq LIIFICXTASISA/YM

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 230:

```
Met Leu Ser Ala Arg Asp Arg Arg Asp Arg His Pro Glu Glu Gly Val
            -55                 -50                 -45

Val Ala Glu Leu Gln Gly Phe Ala Val Asp Lys Ala Phe Leu Thr Ser
        -40                 -35                 -30

His Lys Gly Ile Leu Leu Glu Thr Glu Leu Ala Leu Thr Leu Ile Ile
    -25                 -20                 -15

Phe Ile Cys Xaa Thr Ala Ser Ile Ser Ala Tyr Met Ala Ala Ala Leu
-10                 -5                   1                   5

Leu Glu Phe Phe Ile Thr Leu Ala Phe Leu Phe Leu Tyr Ala Thr Pro
             10                  15                  20

Ala
```

(2) INFORMATION FOR SEQ ID NO: 231:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -13..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.5
seq MLTMSVTLSPLRS/QD

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 231:

```
Met Leu Thr Met Ser Val Thr Leu Ser Pro Leu Arg Ser Gln Asp Leu
            -10                 -5                   1

Asp Pro Met Ala Thr Asp Ala Ser Pro Met Ala Ile Asn Met Thr Pro
    5                   10                  15

Thr Val Glu Gln Gly Leu
 20                 25
```

(2) INFORMATION FOR SEQ ID NO: 232:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -46..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.4
seq FFLLISSVRPISQ/TF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 232:

```
Met Phe Xaa Pro Val Ala Leu Ile Phe Pro Ile Ser Val Ser Asp Pro
```

```
    -45                 -40                 -35

Thr Ile His Pro Ile Thr Gln Ala Gln Asn Leu Glu Ser Xaa Leu Gln
-30                 -25                 -20                 -15

Ser Phe Phe Leu Leu Ile Ser Ser Val Arg Pro Ile Ser Gln Thr Phe
                -10                 -5                  1

Lys Ile Asp Leu Ser Pro Ser Val Arg Ala Thr Gly
        5                   10
```

(2) INFORMATION FOR SEQ ID NO: 233:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -31..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.4
seq WCAVLRSWLAASS/AP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 233:

```
Met Leu Leu Phe Phe Pro Phe Phe Gly Glu Thr Val Ser Leu His His
    -30                 -25                 -20

Pro Cys Trp Cys Ala Val Leu Arg Ser Trp Leu Ala Ala Ser Ser Ala
-15                 -10                 -5                  1

Pro Arg
```

(2) INFORMATION FOR SEQ ID NO: 234:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 56 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Surrenals

(ix) FEATURE:
(A) NAME/KEY: sig_peptide

(B) LOCATION: -35..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.4
seq LVVVCYLSWRVSS/RS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 234:

```
Met Pro Leu Lys Asn Leu Phe Ser Val Gly Leu Trp Asp Pro Tyr Asn
-35                 -30                 -25                 -20

Leu Leu Lys Lys His Val Leu Val Val Val Cys Tyr Leu Ser Trp Arg
                -15                 -10                  -5

Val Ser Ser Arg Ser Trp Thr Leu Leu Ile Thr Pro Val Thr Leu His
              1               5                  10

Ala Ser Leu Ser Thr Gln Ala Arg
     15                  20
```

(2) INFORMATION FOR SEQ ID NO: 235:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 70 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -19..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.4
seq LSHLLPSLRQVIQ/EP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 235:

```
Met Ala Met Ala Gln Lys Leu Ser His Leu Leu Pro Ser Leu Arg Gln
                -15                 -10                  -5

Val Ile Gln Glu Pro Gln Leu Ser Leu Gln Pro Glu Xaa Val Phe Thr
              1               5                  10

Val Asp Arg Ala Glu Val Pro Pro Leu Phe Trp Lys Pro Tyr Ile Tyr
     15                  20                  25

Ala Gly Xaa Arg Pro Leu His Gln Thr Trp Arg Phe Tyr Phe Arg Thr
 30                  35                  40                  45

Leu Phe Gln Gln His Asn
                 50
```

(2) INFORMATION FOR SEQ ID NO: 236:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 74 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -28..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.4
seq LALVALAPHSVQK/SX

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 236:

```
Met Ile Ala Pro Thr Leu Lys Gly Thr Pro Ser Ser Ser Ala Pro Leu
            -25                 -20                 -15

Ala Leu Val Ala Leu Ala Pro His Ser Val Gln Lys Ser Xaa Xaa Phe
        -10                 -5                  1

Pro Pro Asn Leu Leu Thr Ser Pro Pro Ser Val Ala Xaa Ala Glu Ser
  5                 10                  15                  20

Gly Ser Val Ile Thr Leu Ser Ala Xaa Ile Ala Pro Ser Glu Pro Lys
                25                  30                  35

Thr Asn Leu Asn Lys Val Pro Ser Glu Val
            40                  45
```

(2) INFORMATION FOR SEQ ID NO: 237:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -37..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix

(D) OTHER INFORMATION: score 5.3
seq LVESLCLVFNLLS/LP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 237:

```
Met Cys Leu Phe Pro Val Ser Pro Cys Pro Ala Tyr Ser Phe Ser Ser
    -35             -30             -25

Glu Xaa Xaa Gly Ala Val Leu Leu Leu Val Glu Ser Leu Cys Leu Val
-20                 -15                 -10

Phe Asn Leu Leu Ser Leu Pro Pro Arg
 -5                  1
```

(2) INFORMATION FOR SEQ ID NO: 238:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 91 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -15..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.3
seq IAVLFCFLLLIIF/QT

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 238:

```
Met Lys Ile Ala Val Leu Phe Cys Phe Leu Leu Leu Ile Ile Phe Gln
-15                 -10                 -5                   1

Thr Asp Phe Gly Lys Asn Glu Glu Ile Pro Arg Lys Gln Arg Arg Lys
            5                   10                  15

Ile Tyr His Arg Arg Leu Arg Lys Ser Ser Thr Ser His Lys His Arg
        20                  25                  30

Ser Asn Arg Gln Leu Gly Ile Pro Gln Thr Thr Val Phe Thr Pro Val
    35                  40                  45

Ala Arg Leu Pro Ile Val Asn Phe Asp Tyr Ser Met Glu Glu Lys Phe
50                  55                  60                  65

Glu Ser Phe Ser Ser Phe Pro Gly Val Glu Ser
                70                  75
```

(2) INFORMATION FOR SEQ ID NO: 239:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -27..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.2
seq VGAVLLSSLPISP/QY

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 239:

```
Met Cys Ser Pro Arg Ser Pro Leu Asn Leu Ser Leu Val Pro Val Gly
        -25                 -20                 -15

Ala Val Leu Leu Ser Ser Leu Pro Ile Ser Pro Gln Tyr Gly
    -10                 -5                  1
```

(2) INFORMATION FOR SEQ ID NO: 240:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 69 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -48..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.2
seq EVVTLPLTSHCLA/QV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 240:

```
Met Gly Leu His Ile Ser Leu Ile Lys Phe Leu Leu Ala Asn Gly Pro
            -45                 -40                 -35

His Ile Pro Ser His Gln Arg Pro Phe Glu Pro Lys Gly Glu Lys Ser
        -30                 -25                 -20
```

```
Cys Arg Ile Glu Val Val Thr Leu Pro Leu Thr Ser His Cys Leu Ala
    -15                 -10                 -5

Gln Val Ala Ser Ser Asp Leu Ile His Arg Met Arg Thr Ile Thr Gly
  1               5                  10                  15

Thr Ser Ser His Gly
            20
```

(2) INFORMATION FOR SEQ ID NO: 241:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 64 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -26..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.1
seq LLTLYVFVASSMQ/IY

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 241:

```
Met Lys Thr Thr Tyr Val Ile Phe Met Gln Ser Lys Ala Leu Leu Thr
    -25                 -20                 -15

Leu Tyr Val Phe Val Ala Ser Ser Met Gln Ile Tyr Val Leu His Ile
-10                 -5                   1                   5

Ser Asn Tyr Pro Thr Asp Glu His Phe Pro Ile Ile Lys His Phe Tyr
            10                  15                  20

Phe Thr Phe Lys Ile His Phe Ser Lys Ile Ile Tyr Val Gln Tyr Ser
        25                  30                  35
```

(2) INFORMATION FOR SEQ ID NO: 242:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal

(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -15..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.1
seq ALVFLIFLRFINI/SE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 242:

```
Met Asn Ala Leu Val Phe Leu Ile Phe Leu Arg Phe Ile Asn Ile Ser
-15                 -10                 -5                   1

Glu Val Thr Thr Lys Cys Gln Ala Gly
            5                  10
```

(2) INFORMATION FOR SEQ ID NO: 243:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -29..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.1
seq WGFLLTGHSLSHS/SK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 243:

```
Met Gln Leu Gly Pro Leu His Thr Val Ser Thr Pro Phe Phe Phe Cys
                -25                 -20                 -15

Trp Gly Phe Leu Leu Thr Gly His Ser Leu Ser His Ser Ser Lys Ser
            -10                 -5                   1

Cys His Leu
    5
```

(2) INFORMATION FOR SEQ ID NO: 244:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 62 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -57..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.1
seq VGSVCCCVGPLRG/LV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 244:

```
Met Gly Arg Gly Trp Glu Arg Thr Val Cys Ser Leu Gly Trp Arg Gly
        -55                     -50                     -45

Gly Pro Asp Pro Leu Ser Trp Ala Thr Cys Trp Ser Gly Ala Arg Ser
    -40                     -35                     -30

Arg His Thr Arg Val Ser Ser Ile Val Asn Gly Tyr Val Gly Ser Val
-25                     -20                     -15                     -10

Cys Cys Cys Val Gly Pro Leu Arg Gly Leu Val Trp Ala Pro
                -5                      1                   5
```

(2) INFORMATION FOR SEQ ID NO: 245:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -22..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5
seq HLFVTWSSQRALS/HP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 245:

```
Met Pro Glu Ala Val Glu Gln Ser Ala His Leu Phe Val Thr Trp Ser
        -20                     -15                     -10

Ser Gln Arg Ala Leu Ser His Pro Ala Pro Leu
    -5                      1                   5
```

(2) INFORMATION FOR SEQ ID NO: 246:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 49 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -24..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5
seq CWLIALSVPLVFW/VT

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 246:

```
Met Pro Gly Thr His Thr Phe Thr Phe Lys Ser Cys Trp Leu Ile Ala
                -20                 -15                 -10

Leu Ser Val Pro Leu Val Phe Trp Val Thr Phe Trp Pro Cys Asn Phe
             -5                   1               5

Tyr Pro Ser Leu Asp Phe Cys Met Leu Thr Lys Xaa Lys Ser Ile Phe
     10                  15                  20

Ile
 25
```

(2) INFORMATION FOR SEQ ID NO: 247:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -22..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.9
seq LFCLIGLDLLCQV/FS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 247:

```
Met Leu Leu Leu Thr Phe Lys Trp Phe Leu Phe Cys Leu Ile Gly Leu
        -20                 -15                 -10

Asp Leu Leu Cys Gln Val Phe Ser Pro Tyr Phe Leu Ser Glu Lys Val
    -5                  1               5                   10

Ala Asp Leu Leu Phe Tyr Met Ser Leu Phe Phe
                15                  20
```

(2) INFORMATION FOR SEQ ID NO: 248:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 55 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -52..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.9
seq VPNLHLLLPLTTP/QP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 248:

```
Met Ala Thr Thr Gly Arg Arg Gln Ala Glu Pro Pro Pro Val Arg Pro
        -50                 -45                 -40

Ala His Ser Arg Pro Pro Pro Arg Val Pro Gly Ser Ser Ser Leu Gly
    -35                 -30                 -25

Leu Ala Gly Leu Met Ser Pro Val Pro Asn Leu His Leu Leu Leu Pro
-20                 -15                 -10                 -5

Leu Thr Thr Pro Gln Pro Arg
                1
```

(2) INFORMATION FOR SEQ ID NO: 249:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 53 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -16..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.9
seq FIYLQAHFTLCSG/WS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 249:

```
Met Val Pro Phe Ile Tyr Leu Gln Ala His Phe Thr Leu Cys Ser Gly
    -15                 -10                 -5

Trp Ser Ser Thr Tyr Arg Asp Leu Arg Lys Gly Val Tyr Val Pro Tyr
  1               5                  10                  15

Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly Thr Asp Leu Val Ser Ile
            20                  25                  30

Pro His Gly Pro Lys
        35
```

(2) INFORMATION FOR SEQ ID NO: 250:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -14..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.8
seq FFSFLLTINLVSL/QV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 250:

```
Met Phe Phe Ser Phe Leu Leu Thr Ile Asn Leu Val Ser Leu Gln Val
                -10                 -5                   1

Val Ile Leu Asn Arg Val Tyr Leu Asn Gln Pro Asp Ala Arg
        5                   10                  15
```

(2) INFORMATION FOR SEQ ID NO: 251:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 amino acids
(B) TYPE: AMINO ACID

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.8
seq NWGLLCFASECTT/DR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 251:

```
Met Trp Pro Gly Arg Glu Cys Lys Asn Trp Gly Leu Leu Cys Phe Ala
    -20                 -15                 -10

Ser Glu Cys Thr Thr Asp Arg
 -5                  1
```

(2) INFORMATION FOR SEQ ID NO: 252:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -37..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.8
seq CLLLTLRQPPTHS/GI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 252:

```
Met Leu Thr Pro Phe Ser Leu Glu Glu Lys Leu Leu Glu Cys His Tyr
        -35                 -30                 -25

Val Leu Ala Lys Leu Ala Gly Ala Cys Leu Leu Leu Thr Leu Arg Gln
    -20                 -15                 -10

Pro Pro Thr His Ser Gly Ile Pro His Gly
 -5                  1               5
```

(2) INFORMATION FOR SEQ ID NO: 253:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -40..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.7
seq IVFVGLIFLKSSA/HR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 253:

```
Met Lys Arg Ile Lys Ser Met Met Gly Lys Val Glu His Ile Lys Ile
-40                 -35                 -30                 -25

Lys Gly Glu Lys Gln Arg Ser Arg His Val Lys Ile Val Phe Val Gly
                -20                 -15                 -10

Leu Ile Phe Leu Lys Ser Ser Ala His Arg
            -5                   1
```

(2) INFORMATION FOR SEQ ID NO: 254:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -17..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.7
seq LCLFCKICPFTHG/VA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 254:

```
Met Gln Ser Ala Leu Cys Leu Phe Cys Lys Ile Cys Pro Phe Thr His
    -15                 -10                 -5

Gly Val Ala Thr Pro Ala Trp Glu Leu Ser Ser Lys Arg Lys Ala Ser
```

```
 1               5                   10                  15

His Pro Pro Arg
```

(2) INFORMATION FOR SEQ ID NO: 255:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 92 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Surrenals

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -69..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.7
seq SLLCLAFLLGRFL/HM

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 255:

```
Met Met His Asn Ile Ile Val Lys Glu Leu Ile Val Thr Phe Phe Leu
                -65                 -60                 -55

Gly Ile Thr Val Val Gln Met Leu Ile Ser Val Thr Gly Leu Lys Gly
            -50                 -45                 -40

Val Glu Ala Gln Asn Gly Ser Glu Ser Glu Val Phe Val Gly Lys Tyr
        -35                 -30                 -25

Glu Thr Leu Val Phe Tyr Trp Pro Ser Leu Leu Cys Leu Ala Phe Leu
    -20                 -15                 -10

Leu Gly Arg Phe Leu His Met Phe Val Lys Ala Leu Arg Val His Leu
 -5                   1               5                  10

Gly Trp Glu Leu Gln Val Glu Glu Lys Ser Val Leu
            15                  20
```

(2) INFORMATION FOR SEQ ID NO: 256:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 102 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -27..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.6
seq RLLCSRLCQQLRS/KR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 256:

```
Met Ser Asn Cys Leu Gln Asn Phe Leu Lys Ile Thr Ser Thr Arg Leu
        -25                 -20                 -15

Leu Cys Ser Arg Leu Cys Gln Gln Leu Arg Ser Lys Arg Lys Phe Phe
    -10                 -5                  1               5

Gly Thr Val Pro Ile Ser Arg Leu His Arg Arg Val Val Ile Thr Gly
                10                  15                  20

Ile Gly Leu Val Thr Pro Leu Gly Val Gly Thr His Leu Val Trp Asp
            25                  30                  35

Arg Leu Ile Gly Gly Glu Ser Gly Ile Val Ser Leu Val Gly Glu Glu
        40                  45                  50

Tyr Lys Ser Ile Pro Cys Ser Val Ala Ala Tyr Val Pro Arg Gly Ser
    55                  60                  65

Asp Glu Gly Gln Ser Gly
 70                 75
```

(2) INFORMATION FOR SEQ ID NO: 257:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 56 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -16..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.6
seq XGLFLRTTAAARA/CR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 257:

```
Met Ser Gly Xaa Gly Leu Phe Leu Arg Thr Thr Ala Ala Ala Arg Ala
    -15                 -10                 -5

Cys Arg Gly Leu Val Val Ser Thr Ala Asn Arg Arg Leu Leu Arg Thr
```

```
 1              5                   10                  15
Ser Pro Pro Val Arg Ala Phe Ala Lys Glu Leu Phe Leu Gly Lys Ile
            20                  25                  30

Xaa Lys Val Thr Arg Ala Leu Gly
        35                  40
```

(2) INFORMATION FOR SEQ ID NO: 258:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -25..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.5
seq LTWLHLLLSHLKS/SL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 258:

```
Met Asn Pro Leu His Lys His Cys Ala Ala Gly Pro Leu Thr Trp Leu
-25                 -20                 -15                 -10

His Leu Leu Leu Ser His Leu Lys Ser Ser Leu Val Met Pro Pro Lys
                 -5                       1               5
```

(2) INFORMATION FOR SEQ ID NO: 259:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -29..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.5

seq FAVLRVLHLPALT/AP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 259:

```
Met Pro Lys Asp Lys Arg Gly Ala Arg His Asn Ser Pro His Phe Ser
                -25                 -20                 -15

Phe Ala Val Leu Arg Val Leu His Leu Pro Ala Leu Thr Ala Pro Leu
            -10                 -5                  1

Trp Leu Ala Pro Phe Ser Thr Leu Pro Arg
    5                   10
```

(2) INFORMATION FOR SEQ ID NO: 260:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 48 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -39..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.5
seq LCVSRQLLTGART/LF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 260:

```
Met Thr Ile His Val Leu Arg Lys Cys Cys Gln Met Gly Arg Leu Asn
                -35                 -30                 -25

Asn Glu Trp Leu Pro Gly Leu Val Ile Pro Leu Cys Val Ser Arg Gln
            -20                 -15                 -10

Leu Leu Thr Gly Ala Arg Thr Leu Phe Gln Leu Gln Asn Gly Pro Ala
        -5                  1                   5
```

(2) INFORMATION FOR SEQ ID NO: 261:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens

(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -30..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.5
seq IAALLGLLQLRFK/AE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 261:

```
Met Gln Ala Ala Ser Phe Gly Arg Gly Arg Asn Gly Leu Asp Asn Trp
-30                 -25                 -20                 -15

Gly Ile Ala Ala Leu Leu Gly Leu Leu Gln Leu Arg Phe Lys Ala Glu
                -10                 -5                  1
```

(2) INFORMATION FOR SEQ ID NO: 262:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 52 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -43..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.5
seq ENLLLCCHRCTNC/QR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 262:

```
Met Ser Pro Ser Leu Gly Asp Arg Cys Ser Ser Trp Leu His Leu Val
            -40                 -35                 -30

Ser His Leu Glu Ser Ile Ser Gly Pro Leu Leu Asn Ile Pro Glu Asn
        -25                 -20                 -15

Leu Leu Leu Cys Cys His Arg Cys Thr Asn Cys Gln Arg His His Phe
    -10                 -5                  1               5

Cys Ser Val Trp
```

(2) INFORMATION FOR SEQ ID NO: 263:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 47 amino acids
(B) TYPE: AMINO ACID

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -24..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.4
seq YFLLPCLINLAIG/VK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 263:

```
Met Ser Gly Ala Glu Pro Thr Thr Phe Ile Arg Tyr Phe Leu Leu Pro
                -20                 -15                 -10

Cys Leu Ile Asn Leu Ala Ile Gly Val Lys Trp Lys Thr Ala Trp Lys
             -5                   1               5

Arg Gly Glu Arg Gln Leu Asn Asn Thr Val Phe Phe Phe Phe Phe
     10                  15                  20
```

(2) INFORMATION FOR SEQ ID NO: 264:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -22..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.4
seq QLLFSFLLSTIPT/SY

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 264:

```
Met Val Tyr Asp Tyr Phe Ile Ser Gln Gln Leu Leu Phe Ser Phe Leu
        -20                 -15                 -10

Leu Ser Thr Ile Pro Thr Ser Tyr His Leu Ser Leu Thr Cys Gln Arg
     -5                   1               5                  10
```

(2) INFORMATION FOR SEQ ID NO: 265:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -14..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.3
seq LFLCSCSLSLNQL/LT

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 265:

```
Met Leu Phe Leu Cys Ser Cys Ser Leu Ser Leu Asn Gln Leu Leu Thr
                -10                 -5                   1

Tyr Ile Phe Val Val Pro Pro Trp
         5                  10
```

(2) INFORMATION FOR SEQ ID NO: 266:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -15..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.2
seq FLMVLLFRSNKWT/GK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 266:

```
Met Phe Phe Leu Met Val Leu Leu Phe Arg Ser Asn Lys Trp Thr Gly
-15                 -10                 -5                   1

Lys Val Tyr Gly Ala Leu
             5
```

(2) INFORMATION FOR SEQ ID NO: 267:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -29..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.2
seq FLSHVTTSLASSS/AP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 267:

```
Met Leu Pro Leu Gln Gly Leu Cys Thr Cys Tyr Phe Leu His Leu Glu
            -25                 -20                 -15

Phe Leu Ser His Val Thr Thr Ser Leu Ala Ser Ser Ser Ala Pro Ser
        -10                 -5                   1

Pro Lys Pro Ser Val Thr Leu Ser Ser
    5                   10
```

(2) INFORMATION FOR SEQ ID NO: 268:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.2
seq HFFLLLNTILLFG/CA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 268:

```
Met Tyr Phe Tyr Gly Leu Thr Phe His Phe Phe Leu Leu Leu Asn Thr
    -20                 -15                 -10
```

```
Ile Leu Leu Phe Gly Cys Ala Arg
 -5                  1
```

(2) INFORMATION FOR SEQ ID NO: 269:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -29..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.1
seq LWASQGSLQDAQS/ER

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 269:

```
Met Arg Trp Asn Leu Phe Phe Phe Cys Ile Leu Arg Asn Gln Thr Lys
                -25                 -20                 -15

Leu Trp Ala Ser Gln Gly Ser Leu Gln Asp Ala Gln Ser Glu Arg Gly
            -10                  -5                   1

Cys Phe Ser Leu Asn Gln Asp Gly
      5                  10
```

(2) INFORMATION FOR SEQ ID NO: 270:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -14..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.1
seq FIAALFTMAKTWN/QP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 270:

```
Met Phe Ile Ala Ala Leu Phe Thr Met Ala Lys Thr Trp Asn Gln Pro
                -10                 -5                   1

Gly Cys Ser Ser Met Met Gly Trp Ile Lys Lys Met Arg His Met Thr
         5                  10                  15
```

(2) INFORMATION FOR SEQ ID NO: 271:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 61 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -44..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.1
seq LWVXLPXXXVIAS/VV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 271:

```
Met Pro Gly Xaa Lys His Phe Leu Arg Val Phe Arg Xaa Ser Ala Xaa
                -40                 -35                 -30

Arg Ser Val Gly Tyr Xaa Xaa Lys Pro Gly Thr Ser Arg Ala Ser Leu
            -25                 -20                 -15

Trp Val Xaa Leu Pro Xaa Xaa Xaa Val Ile Ala Ser Val Val Thr Phe
        -10                 -5                   1

Ser Xaa His Met Thr Leu Gly Phe Asp Leu Thr Ala Ala
 5                  10                  15
```

(2) INFORMATION FOR SEQ ID NO: 272:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 80 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Surrenals

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -49..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.1
seq VALGPLFVTGHFA/GE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 272:

```
Met Arg Leu Glu Ser Pro Asp Glu Asn Phe Ala Val Val Gln Glu His
                -45                 -40                 -35

Ala Ile His His Ile Asp Gly Pro Leu Arg Arg Phe Leu Leu Leu Glu
            -30                 -25                 -20

Val His Glu Pro Val Ala Leu Gly Pro Leu Phe Val Thr Gly His Phe
        -15                 -10                 -5

Ala Gly Glu Asp Val Ala Glu Arg Arg Glu Asp Val Val Gln Arg Leu
    1               5                   10                  15

Val Val Asp Gly Leu Ala Gln Val Leu Asp Glu Asp Val Ala His Pro
                20                  25                  30
```

(2) INFORMATION FOR SEQ ID NO: 273:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -20..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4
seq EVLLPTVLRGSYC/FS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 273:

```
Met Ala Gly Ser Pro Asp Arg Glu Val Leu Leu Pro Thr Val Leu Arg
-20                 -15                 -10                  -5

Gly Ser Tyr Cys Phe Ser His His Gly
                1               5
```

(2) INFORMATION FOR SEQ ID NO: 274:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 63 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -52..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4
seq RHLFLFEISLVFS/KE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 274:

```
Met His Val Ser Met Leu Glu Gly Phe Asp Glu Asn Leu Asp Val Gln
        -50                 -45                 -40

Gly Glu Leu Ile Leu Gln Asp Ala Phe Gln Val Trp Asp Pro Lys Ser
    -35                 -30                 -25

Leu Ile Arg Lys Gly Arg Glu Arg His Leu Phe Leu Phe Glu Ile Ser
-20                 -15                 -10                 -5

Leu Val Phe Ser Lys Glu Ile Lys Asp Ser Ser Glu His Asn Gly
                1               5                   10
```

(2) INFORMATION FOR SEQ ID NO: 275:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 65 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -39..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.9
seq VLAIGLLHIVLLS/IP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 275:

```
Met Asn Val Gly Thr Ala His Ser Glu Val Asn Xaa Asn Thr Arg Val
                -35                 -30                 -25
```

Met Lys Xaa Arg Gly Ile Trp Leu Ser Tyr Val Leu Ala Ile Gly Leu
-20 -15 -10

Leu His Ile Val Leu Leu Ser Ile Pro Phe Val Ser Val Xaa Val Val
-5 1 5

Trp Thr Leu Thr Asn Leu Ile His Asn Met Gly Met Tyr Ile Phe Leu
10 15 20 25

His

(2) INFORMATION FOR SEQ ID NO: 276:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -16..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.8
seq FVXAIXXYIPTNS/VQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 276:

Met Leu Ser Phe Val Xaa Ala Ile Xaa Xaa Tyr Ile Pro Thr Asn Ser
-15 -10 -5

Val Gln Glu Phe Leu
1 5

(2) INFORMATION FOR SEQ ID NO: 277:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide

(B) LOCATION: -32..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.8
seq TFINITLWLGSLC/QR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 277:

```
Met Asp Glu Tyr Ser Trp Trp Cys His Val Leu Glu Val Val Lys Gly
        -30                 -25                 -20

Gln Met Phe Thr Phe Ile Asn Ile Thr Leu Trp Leu Gly Ser Leu Cys
    -15                 -10                 -5

Gln Arg Phe Phe Tyr Ala Ser Gly
  1                 5
```

(2) INFORMATION FOR SEQ ID NO: 278:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -23..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.8
seq FIFLIQIWKTCLS/FS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 278:

```
Met Arg Arg Lys Gly Gln Gly His Leu Ala Phe Ile Phe Leu Ile Gln
            -20                 -15                 -10

Ile Trp Lys Thr Cys Leu Ser Phe Ser Pro Thr Ser Gly
        -5                  1                   5
```

(2) INFORMATION FOR SEQ ID NO: 279:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens

(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -25..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.8
seq NILFLAVSSFSMP/LP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 279:

```
Met Phe Leu Ile Ser Gly His Val His Leu Ile Tyr Asn Ile Leu Phe
-25                 -20                 -15                 -10

Leu Ala Val Ser Ser Phe Ser Met Pro Leu Pro Cys Leu Tyr Arg
                -5                  1               5
```

(2) INFORMATION FOR SEQ ID NO: 280:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 78 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -64..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.8
seq LFIVVCVICVTLN/FP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 280:

```
Met Thr Pro Arg Ile Leu Ser Glu Val Gln Phe Ser Ala Phe Cys Pro
                -60                 -55                 -50

Tyr Trp Thr Ile Ala Arg Ile Leu Glu Arg Val Gly Ser Ala Cys Phe
            -45                 -40                 -35

Arg Leu Glu Leu Cys Ala Ala Ile Val Gly Tyr Phe Val Leu Asp Val
        -30                 -25                 -20

Arg Thr Phe Leu Phe Ile Val Val Cys Val Ile Cys Val Thr Leu Asn
    -15                 -10                 -5

Phe Pro Arg Xaa Xaa Phe Leu Cys Leu Ser Ser Leu Thr Ala
  1               5                  10
```

(2) INFORMATION FOR SEQ ID NO: 281:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 67 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -14..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq CSLLSGWGQLLRC/VQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 281:

```
Met Cys Ser Leu Leu Ser Gly Trp Gly Gln Leu Leu Arg Cys Val Gln
                -10                  -5                   1

Thr Pro Ala Glu Pro Arg Asp Val Asn Lys Lys Xaa Glu Lys Lys Glu
         5                   10                  15

Lys Tyr Met Pro Leu Val Asp Ser Leu Cys Gly Gly Leu Gly Thr Arg
     20                  25                  30

Asn Ser Asp Cys Leu Arg Gly Gly Ala Gly Arg Gly Arg Asp Gly Arg
 35                  40                  45                  50

Arg Ile Arg
```

(2) INFORMATION FOR SEQ ID NO: 282:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 46 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -37..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq LIPFNFSASGLCA/CS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 282:

Met Leu Phe Ser Phe Cys Phe Pro Val His Phe Trp Asn Pro Ser Ser
-35 -30 -25

Leu Phe Pro Pro Ser Ser Val Ser Leu Ile Pro Phe Asn Phe Ser Ala
-20 -15 -10

Ser Gly Leu Cys Ala Cys Ser Arg Thr Phe Thr His Met Gly
-5 1 5

(2) INFORMATION FOR SEQ ID NO: 283:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -15..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.6
seq WILRILFVIGSXL/EK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 283:

Met Thr Trp Ile Leu Arg Ile Leu Phe Val Ile Gly Ser Xaa Leu Glu
-15 -10 -5 1

Lys Leu Trp Asn Ile Leu Val Ser Tyr Ile Phe
5 10

(2) INFORMATION FOR SEQ ID NO: 284:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 56 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -48..-1

(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.6
seq ILCIFLGLLIIRC/FK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 284:

```
Met Ser Ser Thr Tyr Cys Gly Asn Ser Ser Ala Lys Met Ser Val Asn
            -45                 -40                 -35

Glu Val Ser Ala Phe Ser Leu Ser Leu Glu Gln Lys Thr Gly Phe Ala
        -30                 -25                 -20

Phe Val Gly Ile Leu Cys Ile Phe Leu Gly Leu Leu Ile Ile Arg Cys
    -15                 -10                  -5

Phe Lys Ile Leu Leu Xaa Gln Ser
  1               5
```

(2) INFORMATION FOR SEQ ID NO: 285:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -18..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.6
seq LTMLSMIVGATCY/AM

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 285:

```
Met Thr Asp Ile Trp Leu Thr Met Leu Ser Met Ile Val Gly Ala Thr
            -15                 -10                  -5

Cys Tyr Ala Met Ile Gly
         1
```

(2) INFORMATION FOR SEQ ID NO: 286:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 52 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -17..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.6
seq WIYAFISLGYILG/SG

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 286:

```
Met Xaa Xaa Cys Trp Ile Tyr Ala Phe Ile Ser Leu Gly Tyr Ile Leu
        -15                 -10                 -5

Gly Ser Gly Ile Val Gly Leu Phe Gly Asn Phe Met Phe Lys Leu Leu
     1               5                  10                  15

Arg Asn Cys Gln Thr Val Phe Gln Asp Gly Tyr Ala Ile Leu Pro Phe
                 20                  25                  30

Pro Pro Thr Gly
            35
```

(2) INFORMATION FOR SEQ ID NO: 287:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 46 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -31..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq QLALSWVPPXCRV/SP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 287:

```
Met Phe Ile Arg Thr Leu Lys Thr Thr Val Leu Pro Phe Met Arg Thr
    -30                 -25                 -20

Ala Pro Gln Leu Ala Leu Ser Trp Val Pro Pro Xaa Cys Arg Val Ser
-15                 -10                 -5                   1

Pro Trp Asp Ser Pro Leu Lys Leu Tyr Cys Leu Gln Pro Gln
            5                   10                  15
```

(2) INFORMATION FOR SEQ ID NO: 288:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -24..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq SVLIFCLLPYIYH/FF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 288:

```
Met Arg Thr Gly Ala Glu Met Arg Thr Asn Ser Ser Val Leu Ile Phe
                -20                 -15                 -10

Cys Leu Leu Pro Tyr Ile Tyr His Phe Phe Pro Ala Ser
            -5                   1               5
```

(2) INFORMATION FOR SEQ ID NO: 289:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 55 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -23..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq EGLELGFSHRTFA/FP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 289:

```
Met Ile Val Ile Pro Ser Trp Leu Glu Asn Glu Gly Leu Glu Leu Gly
            -20                 -15                 -10

Phe Ser His Arg Thr Phe Ala Phe Pro Val Thr His Ala Ser Ser Gln
        -5                   1               5
```

```
Tyr Ile Trp Met Asn Xaa Leu Thr Arg Thr Thr Val Ala Ile Ser Val
 10                 15                  20                  25

Tyr Phe Trp Thr His Thr Gly
                30
```

(2) INFORMATION FOR SEQ ID NO: 290:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 48 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -43..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq VISVFLSFLPSYP/GF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 290:

```
Met Leu Lys Lys Glu Ile Ala His His Ser Pro Ser Leu Val Ser Cys
            -40                 -35                 -30

Pro Val Cys Thr Thr Lys Tyr Arg Thr Leu Arg Leu Leu Arg Val Ile
        -25                 -20                 -15

Ser Val Phe Leu Ser Phe Leu Pro Ser Tyr Pro Gly Phe Ser Met Gln
    -10                 -5                   1               5
```

(2) INFORMATION FOR SEQ ID NO: 291:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -30..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix

(D) OTHER INFORMATION: score 3.5
seq CAYSLPGVALTLG/RQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 291:

```
Met Thr Xaa Pro Ser Arg Ala Gln Thr Val Asp Xaa Gly Ile Ala Lys
-30                 -25                 -20                 -15

His Cys Ala Tyr Ser Leu Pro Gly Val Ala Leu Thr Leu Gly Arg Gln
                -10                 -5                   1
```

(2) INFORMATION FOR SEQ ID NO: 292:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9.8
seq LGLLCALLPQHHG/AP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 292:

```
Met Pro Phe Arg Leu Leu Ile Pro Leu Gly Leu Leu Cys Ala Leu Leu
    -20                 -15                 -10

Pro Gln His His Gly Ala Pro Gly Pro Asp Xaa
 -5                  1                   5
```

(2) INFORMATION FOR SEQ ID NO: 293:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -18..-1

(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9.2
seq VFLCSLLAPMVLA/SA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 293:

```
Met Xaa Leu Val Leu Val Phe Leu Cys Ser Leu Leu Ala Pro Met Val
            -15                 -10                 -5

Leu Ala Ser Ala Ala Glu Lys Glu Xaa Xaa Met Xaa Pro Phe His Tyr
         1               5                  10

Asp Tyr Gln Thr Leu
 15
```

(2) INFORMATION FOR SEQ ID NO: 294:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 89 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -30..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9
seq FFLLLLFRGCLIG/AV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 294:

```
Met Ala Leu Arg Arg Pro Pro Arg Leu Arg Leu Cys Ala Arg Leu Pro
-30                 -25                 -20                 -15

Asp Phe Phe Leu Leu Leu Leu Phe Arg Gly Cys Leu Ile Gly Ala Val
                -10                 -5                   1

Asn Leu Lys Ser Ser Asn Arg Thr Pro Val Val Gln Glu Phe Glu Ser
         5                  10                  15

Val Glu Leu Ser Cys Ile Ile Thr Asp Ser Gln Thr Ser Asp Pro Arg
    20                  25                  30

Ile Glu Trp Lys Lys Ile Gln Asp Glu Gln Thr Thr Tyr Val Phe Phe
35                  40                  45                  50

Asp Asn Lys Ile Gln Gly Asp Leu Ala
                55
```

(2) INFORMATION FOR SEQ ID NO: 295:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 113 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -60..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.2
seq VLLTLLLIAFIFL/II

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 295:

```
Met Gly Gly Asn Gly Ser Thr Cys Lys Pro Asp Thr Glu Arg Gln Gly
-60                 -55                 -50                 -45

Thr Leu Ser Thr Ala Ala Pro Thr Thr Ser Pro Ala Pro Cys Leu Ser
                -40                 -35                 -30

Asn His His Asn Lys Lys His Leu Ile Leu Ala Phe Cys Ala Gly Val
            -25                 -20                 -15

Leu Leu Thr Leu Leu Leu Ile Ala Phe Ile Phe Leu Ile Ile Xaa Ser
        -10                 -5                  1

Tyr Arg Lys Tyr His Ser Lys Pro Gln Ala Pro Asp Pro His Ser Asp
  5                  10                  15                  20

Pro Pro Ala Xaa Leu Ser Xaa Ile Pro Gly Glu Ile Thr Tyr Leu Cys
                 25                  30                  35

Gln His Asn Phe Gln Thr Leu Xaa Xaa Lys Arg Ala Ile Thr Trp Leu
            40                  45                  50

Arg
```

(2) INFORMATION FOR SEQ ID NO: 296:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 97 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -39..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.2
seq SALAKLLLTCCSA/LR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 296:

```
Met Gly Thr Ala Asp Ser Asp Glu Met Ala Pro Glu Ala Pro Gln His
                -35                 -30                 -25

Thr His Ile Asp Val His Ile His Gln Glu Ser Ala Leu Ala Lys Leu
            -20                 -15                 -10

Leu Leu Thr Cys Cys Ser Ala Leu Arg Pro Arg Ala Thr Gln Ala Arg
        -5                      1               5

Gly Ser Ser Arg Leu Leu Val Ala Ser Trp Val Met Gln Ile Val Leu
 10                 15                  20                  25

Gly Ile Xaa Ser Ala Val Xaa Gly Gly Phe Phe Tyr Ile Arg Asp Xaa
                30                  35                  40

Thr Leu Xaa Val Thr Ser Gly Ala Ala Ile Trp Thr Gly Ala Val Ala
            45                  50                  55

Val
```

(2) INFORMATION FOR SEQ ID NO: 297:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -14..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.6
seq SLLSFLFARVNLG/SP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 297:

```
Met Ser Leu Leu Ser Phe Leu Phe Ala Arg Val Asn Leu Gly Ser Pro
                -10                 -5                  1

Leu Ser Ala Asn Gly
        5
```

(2) INFORMATION FOR SEQ ID NO: 298:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 80 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -43..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.2
seq WWCCPARLTLTSG/WP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 298:

```
Met Ala Arg Ser Pro Leu Arg Arg Arg Gly Arg Pro Thr Trp Ser Leu
            -40                 -35                 -30

Ser Thr Pro Arg Pro Gly Ser Pro Thr Ser Ser Ser Arg Ser Trp Trp
        -25                 -20                 -15

Cys Cys Pro Ala Arg Leu Thr Leu Thr Ser Gly Trp Pro Ala Thr Pro
    -10                 -5                  1                5

Arg Arg Phe Ser Thr Thr Ser Thr Cys Ser Ile Ala Pro Ser Arg Ser
                 10                  15                  20

Ser Ala Gln Glu Arg Arg Val Arg Arg Trp Pro Cys Ala Thr His Ser
            25                  30                  35
```

(2) INFORMATION FOR SEQ ID NO: 299:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -19..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7

seq TLLLACHLQLEVG/VV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 299:

```
Met Lys Ile Ile Thr Thr Thr Leu Leu Leu Ala Cys His Leu Gln Leu
                -15                 -10                 -5

Glu Val Gly Val Val Val Gly Gly Glu Val Asp Met Ala Thr Leu Gln
            1               5                   10

Ile Thr Thr Ala Ser
    15
```

(2) INFORMATION FOR SEQ ID NO: 300:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -37..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq FLGVLALLGYLAV/RP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 300:

```
Met Leu Met Pro Val Val Gly Arg Gly Asn Gly Ile Pro Gln Thr Val
        -35                 -30                 -25

Ser Glu Trp Leu Arg Leu Leu Pro Phe Leu Gly Val Leu Ala Leu Leu
    -20                 -15                 -10

Gly Tyr Leu Ala Val Arg Pro Gly
-5                      1
```

(2) INFORMATION FOR SEQ ID NO: 301:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 79 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens

(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -49..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.6
seq PPFFLCLQCFTRG/FE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 301:

```
Met Asp Arg Leu Gly Ser Phe Ser Asn Asp Pro Ser Asp Lys Pro Pro
                -45                 -40                 -35

Cys Arg Gly Cys Ser Ser Tyr Leu Met Glu Pro Tyr Ile Lys Cys Ala
            -30                 -25                 -20

Glu Cys Gly Pro Pro Pro Phe Phe Leu Cys Leu Gln Cys Phe Thr Arg
        -15                 -10                 -5

Gly Phe Glu Tyr Lys Lys His Gln Ser Asp His Thr Tyr Glu Ile Met
    1               5                   10                  15

Ala Gly Cys Ser Gln Ser Asn Val His Gln Asp Gln Gly Gly Gln
                20                  25                  30
```

(2) INFORMATION FOR SEQ ID NO: 302:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 111 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -87..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.6
seq GLLLLYMVYLTLV/EP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 302:

```
Met Ser Asp Val Asn Val Ser Ala Leu Pro Ile Lys Lys Asn Ser Gly
        -85                 -80                 -75

His Ile Tyr Asn Lys Asn Ile Ser Gln Lys Asp Cys Asp Cys Leu His
    -70                 -65                 -60

Val Val Glu Pro Met Pro Val Arg Gly Pro Asp Val Glu Ala Tyr Cys
-55                 -50                 -45                 -40
```

```
Leu Arg Cys Glu Cys Lys Tyr Glu Glu Arg Ser Ser Val Thr Ile Lys
                -35                 -30                 -25

Val Thr Ile Ile Ile Tyr Leu Ser Ile Leu Gly Leu Leu Leu Leu Tyr
            -20                 -15                 -10

Met Val Tyr Leu Thr Leu Val Glu Pro Ile Leu Xaa Arg Arg Leu Phe
         -5                   1                   5

Gly His Ala Gln Leu Ile Gln Ser Asp Asp Xaa Ile Gly Gly Leu
 10                  15                  20
```

(2) INFORMATION FOR SEQ ID NO: 303:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 48 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -46..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq ALSLSLSMAPPNP/GP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 303:

```
Met Ala Xaa Cys Arg Arg Cys Arg Ser Gln Arg Arg Ser His Cys Cys
    -45                 -40                 -35

Gln Asp Arg Arg Leu Arg Arg Pro Arg Leu Thr Leu Trp Arg His His
-30                 -25                 -20                 -15

Thr Ala Leu Ser Leu Ser Leu Ser Met Ala Pro Pro Asn Pro Gly Pro
                -10                  -5                   1
```

(2) INFORMATION FOR SEQ ID NO: 304:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 96 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -36..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq LALTALSVXRKXS/XX

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 304:

```
Met Thr Arg Leu Gly Gly Lys Gly Gly Gln Gln Phe Pro Pro Gly Gln
    -35                 -30                 -25

Lys Ile Ile Ser Lys Asp Ile Leu Ala Leu Thr Ala Leu Ser Val Xaa
-20                 -15                 -10                  -5

Arg Lys Xaa Ser Xaa Xaa Xaa Xaa Xaa Xaa Thr Ser Lys Glu Thr Xaa
                      1               5                  10

Asp Asn Gln Asp Ser Val Lys Glu Asn Arg Glu Lys Asp Leu Leu Asp
         15                  20                  25

Ile Ile Lys Gly Thr Lys Val Glu Leu Ser Thr Val Asn Val Gln Thr
     30                  35                  40

Thr Lys Pro Pro Asn Arg Ser Ser Leu Lys Ser Tyr Asn Trp Arg Ala
 45                  50                  55                  60
```

(2) INFORMATION FOR SEQ ID NO: 305:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -20..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 12.4
seq ALLLGALLGTAWA/RR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 305:

```
Met Lys Gly Trp Gly Trp Leu Ala Leu Leu Leu Gly Ala Leu Leu Gly
-20                 -15                 -10                  -5

Thr Ala Trp Ala Arg Arg Ser Arg
                  1
```

(2) INFORMATION FOR SEQ ID NO: 306:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -16..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 11.4
seq LLCLLLLFGGGDP/RR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 306:

```
Met His Arg Leu Leu Cys Leu Leu Leu Leu Phe Gly Gly Gly Asp Pro
    -15                 -10                 -5

Arg Arg Arg Ala Glu Ile Arg Leu Gln Ala Thr Ile Cys Ser Arg Pro
  1               5                  10                  15

Leu Arg Lys Thr Thr Ser Gly Arg Gly Gly Pro Pro Trp
            20                  25
```

(2) INFORMATION FOR SEQ ID NO: 307:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 128 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -22..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 11.1
seq QLLALFFLPFCLC/QD

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 307:

```
Met Leu Trp Arg Gln Leu Ile Tyr Trp Gln Leu Leu Ala Leu Phe Phe
        -20                 -15                 -10
```

```
Leu Pro Phe Cys Leu Cys Gln Asp Glu Tyr Met Glu Ser Pro Gln Thr
     -5                  1               5                  10

Gly Gly Leu Pro Pro Asp Cys Ser Lys Cys Cys His Gly Asp Tyr Ser
                15                  20                  25

Phe Arg Gly Tyr Gln Gly Pro Pro Gly Pro Pro Gly Pro Pro Gly Ile
            30                  35                  40

Pro Gly Asn His Gly Asn Asn Gly Asn Asn Gly Ala Thr Gly His Glu
        45                  50                  55

Gly Ala Lys Gly Glu Lys Gly Asp Lys Gly Asp Leu Gly Pro Arg Gly
    60                  65                  70

Glu Arg Gly Gln His Gly Pro Lys Gly Glu Lys Gly Tyr Pro Gly Ile
 75                  80                  85                  90

Pro Pro Glu Leu Gln Ile Ala Phe Met Ala Ser Leu Xaa Pro Thr Ser
                95                 100                 105
```

(2) INFORMATION FOR SEQ ID NO: 308:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -17..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9.2
seq LLLLVAASAMVRS/XA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 308:

```
Met Arg Leu Leu Leu Leu Leu Leu Val Ala Ala Ser Ala Met Val Arg
        -15                 -10                     -5

Ser Xaa Ala Ser Ala Asn Leu Gly Gly Val Pro Ser Lys Arg Leu Lys
     1               5                  10                  15

Met Gln Tyr Thr Thr
                20
```

(2) INFORMATION FOR SEQ ID NO: 309:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 132 amino acids

(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -25..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9
seq ALLVLLGVAASLC/VR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 309:

```
Met Ser Ser Gly Xaa Glu Leu Leu Trp Pro Gly Ala Ala Leu Leu Val
-25                 -20                 -15                 -10

Leu Leu Gly Val Ala Ala Ser Leu Cys Val Arg Cys Ser Arg Pro Gly
                 -5                   1               5

Ala Lys Arg Ser Glu Lys Ile Tyr Gln Gln Arg Ser Leu Arg Glu Asp
         10                  15                  20

Gln Gln Ser Phe Thr Gly Ser Arg Thr Tyr Ser Leu Val Gly Gln Ala
     25                  30                  35

Trp Pro Gly Pro Leu Ala Asp Met Ala Pro Thr Arg Lys Asp Lys Leu
 40                  45                  50                  55

Leu Gln Phe Xaa Pro Ser Leu Glu Xaa Pro Ser Ile Phe Gln Xaa Xaa
                 60                  65                  70

Glu Xaa Gln Pro Val Cys Val Cys Ala Ala His Ala Gln Val Gln Xaa
             75                  80                  85

Xaa Gln Arg Lys Ser Thr Ser Arg Glu Val Cys Val Arg Thr Asn Arg
         90                  95                 100

Ala Leu Arg Gly
    105
```

(2) INFORMATION FOR SEQ ID NO: 310:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 125 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -78..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.8
seq SCLGLTLMPFASS/FP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 310:

```
Met Thr Lys Glu Ile Phe Phe Phe Thr Val Glu Leu Val Cys Glu Asn
            -75                 -70                 -65

Lys Glu Leu Cys Ser Ser Pro Arg Trp Arg Asn Ala Ile Gln Lys Ser
        -60                 -55                 -50

Asn Phe Ser Lys Val Thr Ser Phe Phe Met Ser Cys His His Phe Lys
    -45                 -40                 -35

Gly Leu Ala Pro Leu Pro His Val Tyr Thr Gln Gly Asn Cys Arg Pro
-30                 -25                 -20                 -15

Ile Ser Cys Leu Gly Leu Thr Leu Met Pro Phe Ala Ser Ser Phe Pro
                -10                  -5                   1

Glu Val Lys Val Pro Val Met Tyr Ser His Arg Asn Ile Phe Gln Leu
         5                  10                  15

Phe Met Ser Phe Thr Thr Lys Lys Lys Ile Gln Ser Gly Trp Ser Thr
    20                  25                  30

Thr Leu Ser Ile Phe Leu Val Arg Asn Phe Leu Leu Ile
35                  40                  45
```

(2) INFORMATION FOR SEQ ID NO: 311:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.5
seq YFRALCLPRGAWG/FP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 311:

```
Met Thr Thr Asp Ile Gly Cys Leu Tyr Phe Arg Ala Leu Cys Leu Pro
```

```
    -20                  -15                 -10

Arg Gly Ala Trp Gly Phe Pro Ser Leu Gln Ile Lys Gly
 -5                  1               5
```

(2) INFORMATION FOR SEQ ID NO: 312:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -22..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.4
seq LTCLFLFLNLRWS/RH

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 312:

```
Met Val Pro Ser Leu Val Ile Pro Asp Leu Thr Cys Leu Phe Leu Phe
        -20                 -15                 -10

Leu Asn Leu Arg Trp Ser Arg His Val
     -5                  1
```

(2) INFORMATION FOR SEQ ID NO: 313:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 56 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -15..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7
seq LRLLKLAATSASA/RV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 313:

Met Ala Leu Arg Leu Leu Lys Leu Ala Ala Thr Ser Ala Ser Ala Arg
-15 -10 -5 1

Val Val Xaa Ala Xaa Ala Gln Arg Val Arg Gly Ile His Ser Ser Val
5 10 15

Gln Cys Lys Leu Arg Tyr Gly Met Trp His Phe Leu Leu Gly Asp Lys
20 25 30

Ala Ser Lys Arg Leu Thr Val Gln
35 40

(2) INFORMATION FOR SEQ ID NO: 314:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -20..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.9
seq VLLFLYSVLLTKG/IE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 314:

Met Trp Gly Asn Lys Phe Gly Val Leu Leu Phe Leu Tyr Ser Val Leu
-20 -15 -10 -5

Leu Thr Lys Gly Ile Glu Asn Ile Lys Asn Glu Ile Glu Asp Ala Ser
1 5 10

Glu Pro Leu Ile Asp Pro Val Tyr Gly His Gly Xaa
15 20

(2) INFORMATION FOR SEQ ID NO: 315:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -27..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.8
seq VFVCSSVLGQSWG/GF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 315:

```
Met Tyr Thr Phe Arg Lys Leu Ser Pro Tyr Leu Asn Lys Ile Val Phe
        -25                 -20                 -15

Val Cys Ser Ser Val Leu Gly Gln Ser Trp Gly Gly Phe Phe Ser Asn
    -10                 -5                  1               5

Leu Ser Glu Thr Leu Ser Ala Thr Leu Phe Asn Gly
                10                  15
```

(2) INFORMATION FOR SEQ ID NO: 316:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.4
seq TFCLIFGLGAVWG/LV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 316:

```
Met Glu Ser Arg Val Leu Leu Arg Thr Phe Cys Leu Ile Phe Gly Leu
    -20                 -15                 -10

Gly Ala Val Trp Gly Leu Val
 -5                  1
```

(2) INFORMATION FOR SEQ ID NO: 317:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 62 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -37..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.4
seq ILIFLGFFLGLFH/QF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 317:

```
Met Leu Val Leu Lys Lys His Ser Val Asn Ile Ala Ala Gln Thr Cys
        -35                 -30                 -25

Phe Lys Phe Asn Phe Ile Phe Arg Ile Leu Ile Phe Leu Gly Phe Phe
    -20                 -15                 -10

Leu Gly Leu Phe His Gln Phe Leu Phe Leu Phe Leu Phe Ala Gly Asn
 -5                   1               5                  10

Leu Ser Ser Tyr Leu Leu Lys Gln Ser Lys Ile Gln Ala Arg
             15                  20                  25
```

(2) INFORMATION FOR SEQ ID NO: 318:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -32..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.3
seq TVVLCVGCSTVLC/QP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 318:

```
Met Asp Val Lys Cys Pro Gly Cys Tyr Lys Ile Thr Thr Val Phe Ser
        -30                 -25                 -20

His Ala Gln Thr Val Val Leu Cys Val Gly Cys Ser Thr Val Leu Cys
    -15                 -10                  -5

Gln Pro Thr Gly Gly Lys Ala Arg Leu Thr Glu Gly Cys Ser Phe Arg
```

1 5 10 15

Arg Lys

(2) INFORMATION FOR SEQ ID NO: 319:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -16..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.2
seq ILSVLHALPAGIA/WS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 319:

Met Cys Ile Ile Leu Ser Val Leu His Ala Leu Pro Ala Gly Ile Ala
-15 -10 -5

Trp Ser Arg Glu Lys Gly
1 5

(2) INFORMATION FOR SEQ ID NO: 320:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 70 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Surrenals

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -60..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.9
seq TWLLLGALEPASE/RP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 320:

```
Met Leu Val Val Glu Ala Ser Ser Ser Val Arg Leu Ala Ser Ser Glu
-60                 -55                 -50                 -45

Val Thr Ser Trp Ser Ile Leu Val Thr Pro Ser Ala Ser Thr Pro Ile
                -40                 -35                 -30

Ile Ser Leu Ser Ala Gly Pro Leu Arg Thr Pro Ser His Ser Lys Thr
            -25                 -20                 -15

Trp Leu Leu Leu Gly Ala Leu Glu Pro Ala Ser Glu Arg Pro Cys Ser
        -10                 -5                  1

Ser Val Leu Arg Ser Arg
  5                 10
```

(2) INFORMATION FOR SEQ ID NO: 321:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 35 amino acids
- (B) TYPE: AMINO ACID
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
- (A) ORGANISM: Homo Sapiens
- (D) DEVELOPMENTAL STAGE: Fetal
- (F) TISSUE TYPE: brain

(ix) FEATURE:
- (A) NAME/KEY: sig_peptide
- (B) LOCATION: -27..-1
- (C) IDENTIFICATION METHOD: Von Heijne matrix
- (D) OTHER INFORMATION: score 5.7
  seq LSLQLIAFPTVSC/EI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 321:

```
Met Tyr Ser Phe Pro Thr Thr Val Val Glu Glu Ile Leu Ser Leu Ser
        -25                 -20                 -15

Leu Gln Leu Ile Ala Phe Pro Thr Val Ser Cys Glu Ile Leu Leu Glu
    -10                 -5                  1               5

Ile Thr Arg
```

(2) INFORMATION FOR SEQ ID NO: 322:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 amino acids
- (B) TYPE: AMINO ACID
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -16..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.6
seq LLPLRSLLALVRE/SR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 322:

```
Met Leu Met Leu Leu Pro Leu Arg Ser Leu Leu Ala Leu Val Arg Glu
    -15                 -10                 -5

Ser Arg Ala Arg
  1
```

(2) INFORMATION FOR SEQ ID NO: 323:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 103 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cerebellum

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -25..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.5
seq AQLFACLLRLGTQ/QV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 323:

```
Met Val Pro Leu Val Ala Val Val Ser Gly Pro Arg Ala Gln Leu Phe
-25                 -20                 -15                 -10

Ala Cys Leu Leu Arg Leu Gly Thr Gln Gln Val Gly Pro Leu Gln Leu
                -5                   1                   5

His Thr Gly Ala Ser His Ala Ala Arg Asn His Tyr Glu Val Leu Val
        10                  15                  20

Leu Gly Gly Gly Ser Gly Gly Ile Thr Met Ala Ala Arg Met Lys Arg
    25                  30                  35

Lys Val Gly Ala Glu Asn Val Ala Ile Val Glu Pro Ser Glu Arg His
 40                  45                  50                  55

Phe Tyr Gln Pro Ile Trp Thr Leu Val Gly Ala Gly Ala Xaa Asn Cys
```

60 65 70

Pro His Leu Val Val Pro Arg
75

(2) INFORMATION FOR SEQ ID NO: 324:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -20..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.4
seq AFVIACVLSLIST/IY

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 324:

Met Asp Asn Arg Phe Ala Thr Ala Phe Val Ile Ala Cys Val Leu Ser
-20 -15 -10 -5

Leu Ile Ser Thr Ile Tyr Met Ala Arg
1 5

(2) INFORMATION FOR SEQ ID NO: 325:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 120 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -42..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.4
seq WTLLLTSLDGHLL/EP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 325:

```
Met Pro Glu Tyr Cys Gly Asn Glu Val Thr Pro Thr Glu Ala Ala Gln
        -40                 -35                 -30

Ala Pro Glu Val Thr Tyr Glu Ala Glu Glu Gly Ser Leu Trp Thr Leu
    -25                 -20                 -15

Leu Leu Thr Ser Leu Asp Gly His Leu Leu Glu Pro Asp Ala Glu Tyr
-10                 -5                  1               5

Leu His Trp Leu Leu Thr Asn Ile Pro Gly Asn Arg Val Ala Glu Gly
            10                  15                  20

Gln Val Thr Cys Pro Tyr Leu Pro Pro Phe Pro Ala Arg Gly Ser Gly
        25                  30                  35

Ile His Arg Leu Ala Phe Leu Leu Phe Lys Gln Asp Gln Pro Ile Asp
    40                  45                  50

Phe Ser Glu Asp Ala Arg Pro Ser Pro Cys Tyr Gln Leu Xaa Gln Arg
55                  60                  65                      70

Thr Phe Arg Thr Phe Asp Phe Tyr
                75
```

(2) INFORMATION FOR SEQ ID NO: 326:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Surrenals

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -15..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.4
seq RVLCAPAAGAVRA/LR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 326:

```
Met Asn Arg Val Leu Cys Ala Pro Ala Ala Gly Ala Val Arg Ala Leu
-15                 -10                 -5                      1

Arg Leu Ile Gly Trp Ala Ser Arg Ser Leu His Pro Leu Pro Gly Lys
            5                   10                  15
```

(2) INFORMATION FOR SEQ ID NO: 327:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 amino acids

(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -23..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.3
seq MLALLLTAALIFF/AI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 327:

```
Met Ala Phe Thr Phe Ala Ala Phe Cys Tyr Met Leu Ala Leu Leu Leu
            -20                 -15                 -10

Thr Ala Ala Leu Ile Phe Phe Ala Ile Trp His Ile Ile Ala Phe Asp
        -5                   1               5

Glu Leu Lys Thr Asp Tyr Lys Asn Pro Ile Asp Gln Leu
 10                  15                  20
```

(2) INFORMATION FOR SEQ ID NO: 328:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -17..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.3
seq XEXLLAFHHDCEA/SP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 328:

```
Met Xaa Xaa Xaa Xaa Glu Xaa Leu Leu Ala Phe His His Asp Cys Glu
        -15                 -10                      -5

Ala Ser Pro Ala Thr Trp Asn Leu Ser Pro Arg
     1               5                   10
```

(2) INFORMATION FOR SEQ ID NO: 329:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 76 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -36..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.2
seq PLRLLNLLILIEG/SV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 329:

```
Met Gly Pro Tyr Asn Val Ala Val Pro Ser Asp Val Ser His Ala Arg
    -35                 -30                 -25

Phe Tyr Phe Leu Phe His Arg Pro Leu Arg Leu Leu Asn Leu Leu Ile
-20                 -15                 -10                  -5

Leu Ile Glu Gly Ser Val Val Phe Tyr Gln Leu Tyr Ser Leu Leu Arg
                 1               5                  10

Ser Glu Lys Trp Asn His Thr Leu Ser Met Ala Leu Ile Leu Phe Cys
        15                  20                  25

Asn Tyr Tyr Val Leu Phe Lys Leu Leu Arg Asp Gln
     30                  35                  40
```

(2) INFORMATION FOR SEQ ID NO: 330:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 47 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -23..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.2

seq IGVGLYLLASAAA/FY

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 330:

```
Met Met Asn Phe Arg Gln Arg Met Gly Trp Ile Gly Val Gly Leu Tyr
            -20                 -15                 -10

Leu Leu Ala Ser Ala Ala Ala Phe Tyr Tyr Val Phe Glu Ile Ser Glu
        -5                  1               5

Thr Tyr Asn Arg Leu Ala Leu Glu His Ile Gln Gln His Xaa Gly
 10                 15                  20
```

(2) INFORMATION FOR SEQ ID NO: 331:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 130 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -118..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.1
seq AFXVVCWLGPCEA/MH

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 331:

```
Met Leu Phe Ala Ser Gly Gly Phe Xaa Val Lys Leu Tyr Asp Ile Glu
            -115                -110                -105

Gln Gln Gln Ile Arg Asn Ala Leu Glu Asn Ile Arg Lys Glu Met Lys
        -100                -95                 -90

Leu Leu Glu Gln Ala Gly Ser Leu Lys Gly Ser Leu Ser Val Glu Glu
    -85                 -80                 -75

Gln Leu Ser Leu Ile Ser Gly Cys Pro Asn Ile Gln Glu Ala Val Glu
-70                 -65                 -60                 -55

Gly Ala Met His Ile Gln Glu Cys Val Pro Glu Asp Leu Glu Leu Lys
                -50                 -45                 -40

Lys Lys Ile Phe Ala Gln Leu Asp Ser Ile Ile Asp Glu Ser Ser Asp
            -35                 -30                 -25

Leu Lys Arg Phe Xaa Phe Leu Ser His Ala Phe Xaa Val Val Cys Trp
        -20                 -15                 -10

Leu Gly Pro Cys Glu Ala Met His Arg Gly Ser Ser Cys Glu Ser Ala
```

-5 1 5 10

Ile Leu

(2) INFORMATION FOR SEQ ID NO: 332:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -22..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.1
seq LCSLPLSPSAVCP/AP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 332:

Met Gln Cys Phe Leu Gly Gly Leu Gly Leu Cys Ser Leu Pro Leu Ser
-20 -15 -10

Pro Ser Ala Val Cys Pro Ala Pro Thr Ser Ala Pro Trp Trp Glu Gly
-5 1 5 10

Ala Leu

(2) INFORMATION FOR SEQ ID NO: 333:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 63 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -13..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.9
seq MSSFLLSFSQSLS/NV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 333:

```
Met Ser Ser Phe Leu Leu Ser Phe Ser Gln Ser Leu Ser Asn Val Pro
            -10                 -5                   1

Ser Ala Leu Gln Xaa Pro Gln Ile Thr Phe Phe Gln His Pro Leu Ser
     5                  10                  15

Ser Val Met Pro Val Trp Thr Cys Ser Val Val Pro Cys Asp Lys Thr
 20                 25                  30                  35

Xaa Gln Tyr Ser Tyr Cys Phe Tyr Cys Val Leu Gly Thr Val Lys
                40                  45                  50
```

(2) INFORMATION FOR SEQ ID NO: 334:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 76 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -13..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.9
seq MLTASLAFQLVDG/VS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 334:

```
Met Leu Thr Ala Ser Leu Ala Phe Gln Leu Val Asp Gly Val Ser Trp
            -10                 -5                   1

Asn Phe Ser Val Ser Lys Met Leu Ala Ser Pro Ser Thr Ser Gly Gln
     5                  10                  15

Leu Ser Gln Phe Gly Ala Ser Leu Tyr Gly Gln Gln Ser Ala Leu Gly
 20                 25                  30                  35

Leu Pro Met Arg Gly Met Ser Asn Asn Thr Pro Gln Leu Asn Arg Ser
                40                  45                  50

Leu Ser Gln Xaa Leu Ser Tyr Arg Ala Thr Ser Arg
            55                  60
```

(2) INFORMATION FOR SEQ ID NO: 335:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -25..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.9
seq LSAFNFLVCLSLG/RG

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 335:

```
Met Tyr Xaa Arg Arg Glu Leu Ser Ile Leu Cys Ile Leu Ser Ala Phe
-25                 -20                 -15                 -10

Asn Phe Leu Val Cys Leu Ser Leu Gly Arg Gly
                -5                  1
```

(2) INFORMATION FOR SEQ ID NO: 336:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -22..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.8
seq IVFGVSWVMLVYS/AS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 336:

```
Met Gly Leu Ser Ala Met Asp Thr Ser Ile Val Phe Gly Val Ser Trp
    -20                 -15                 -10

Val Met Leu Val Tyr Ser Ala Ser Phe Arg Arg Cys Xaa
    -5                  1               5
```

(2) INFORMATION FOR SEQ ID NO: 337:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 69 amino acids
(B) TYPE: AMINO ACID

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -65..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.8
seq IILFSAIVGFIYG/YV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 337:

```
Met Tyr Phe Trp Arg Asp Val Ala Val Ser Leu Asp Thr Leu Trp Ala
-65                 -60                 -55                 -50

Leu Pro Arg Gln Gln Pro Gly Leu Gly Asn Asn Arg Val Leu Gly Leu
                -45                 -40                 -35

Leu Ser Gly Thr Asn Lys Asp Tyr Lys Gly Gln Lys Leu Ala Glu Gln
            -30                 -25                 -20

Met Phe Gln Gly Ile Ile Leu Phe Ser Ala Ile Val Gly Phe Ile Tyr
        -15                 -10                 -5

Gly Tyr Val Ala Ala
    1
```

(2) INFORMATION FOR SEQ ID NO: 338:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 54 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -26..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.7
seq LLICXLXIGTATP/VR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 338:

```
Met His Trp Gly Lys Arg Trp Xaa Leu Xaa Xaa Gly Gly Leu Leu Ile
    -25                 -20                 -15
```

```
Cys Xaa Leu Xaa Ile Gly Thr Ala Thr Pro Val Arg Xaa Pro Asn Gly
-10                 -5                  1               5

Arg Gln Val Leu Val Pro Xaa Gly Tyr Pro Arg Pro Gly Leu Gly Ala
            10                  15                  20

Val Gly Cys Gly Glu Ala
        25
```

(2) INFORMATION FOR SEQ ID NO: 339:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -29..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.6
seq ALGLXTCLSVLFG/YA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 339:

```
Met Ala Xaa Arg Tyr Asn Arg Leu Thr Val Leu Ala Gly Ala Xaa Leu
                -25                 -20                 -15

Ala Leu Gly Leu Xaa Thr Cys Leu Ser Val Leu Phe Gly Tyr Ala Pro
            -10                 -5                  1

Gln Ser Ser Pro
    5
```

(2) INFORMATION FOR SEQ ID NO: 340:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -29..-1

(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.5
seq FMTCILCRPPISS/CV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 340:

```
Met Gly Phe Thr Gly Phe Phe Thr Ala Thr Cys Phe Ile Ser Lys Val
                -25                 -20                 -15

Phe Met Thr Cys Ile Leu Cys Arg Pro Pro Ile Ser Ser Cys Val Leu
            -10                 -5                  1

Glu Cys Gly
    5
```

(2) INFORMATION FOR SEQ ID NO: 341:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -35..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.4
seq LIVLLPVLFFSLK/NF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 341:

```
Met Ile Met Tyr Leu Phe Val Ile Cys Val Ile Phe Glu Ile Ile Arg
-35                 -30                 -25                 -20

Asn Tyr Ala Phe Ser Ile Leu Ile Val Leu Leu Pro Val Leu Phe Phe
                -15                 -10                 -5

Ser Leu Lys Asn Phe Ile Leu Ser Thr Gln
            1                   5
```

(2) INFORMATION FOR SEQ ID NO: 342:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 58 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -37..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.3
seq LWEKLTLLSPGIA/VT

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 342:

```
Met Ser Thr Val Gly Leu Xaa His Phe Pro Xaa Pro Leu Thr Arg Ile
    -35                 -30                 -25

Cys Pro Ala Pro Trp Gly Leu Arg Leu Trp Glu Lys Leu Thr Leu Leu
    -20                 -15                 -10

Ser Pro Gly Ile Ala Val Thr Pro Val Gln Met Ala Gly Lys Lys Asp
 -5                     1               5                   10

Xaa Pro Ala Leu Leu Ser Leu Asp Glu Asn
            15                  20
```

(2) INFORMATION FOR SEQ ID NO: 343:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 79 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -13..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.3
seq MLALAXHLSTVES/EK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 343:

```
Met Leu Ala Leu Ala Xaa His Leu Ser Thr Val Glu Ser Glu Lys Gln
            -10                 -5                  1

Lys Leu Arg Ala Gln Val Arg Arg Leu Cys Gln Glu Asn Gln Trp Leu
    5                   10                  15

Arg Asp Glu Leu Ala Gly Thr Gln Gln Arg Leu Gln Arg Ser Glu Gln
20                  25                  30                  35
```

```
Ala Val Ala Gln Leu Glu Glu Glu Lys Lys His Leu Glu Phe Leu Gly
                 40                  45                  50

Gln Leu Arg Gln Tyr Asp Glu Asp Gly His Thr Ser Glu Ala Gly
            55                  60                  65
```

(2) INFORMATION FOR SEQ ID NO: 344:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 53 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -34..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.3
seq QLFAFLNLLPVEA/DI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 344:

```
Met Leu Leu Ser Ile Gly Met Leu Met Leu Xaa Ala Thr Gln Val Tyr
                -30                 -25                 -20

Thr Ile Leu Thr Val Gln Leu Phe Ala Phe Leu Asn Leu Leu Pro Val
            -15                 -10                 -5

Glu Ala Asp Ile Leu Ala Tyr Asn Phe Glu Asn Ala Ser Gln Thr Phe
         1               5                  10

Asp Asp Leu Pro Ala
 15
```

(2) INFORMATION FOR SEQ ID NO: 345:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 76 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:

(A) NAME/KEY: sig_peptide
(B) LOCATION: -37..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.3
seq LWEKLTLLSPGIA/VT

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 345:

```
Met Ser Thr Val Gly Leu Phe His Phe Pro Thr Pro Leu Thr Arg Ile
        -35                 -30                 -25

Cys Pro Ala Pro Trp Gly Leu Arg Leu Trp Glu Lys Leu Thr Leu Leu
    -20                 -15                 -10

Ser Pro Gly Ile Ala Val Thr Pro Val Gln Met Ala Gly Lys Lys Asp
 -5                  1               5                  10

Tyr Pro Ala Leu Leu Ser Leu Asp Glu Xaa Glu Leu Glu Glu Gln Phe
            15                  20                  25

Val Lys Gly His Gly Pro Gly Gly Gln Ala Thr Arg
        30                  35
```

(2) INFORMATION FOR SEQ ID NO: 346:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 165 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -33..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.2
seq LLNFLGLWSWICK/KW

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 346:

```
Met Glu Leu Thr Ile Phe Ile Leu Arg Leu Ala Ile Tyr Ile Leu Thr
            -30                 -25                 -20

Phe Pro Leu Tyr Leu Leu Asn Phe Leu Gly Leu Trp Ser Trp Ile Cys
        -15                 -10                  -5

Lys Lys Trp Phe Pro Tyr Phe Leu Val Arg Phe Thr Val Ile Tyr Asn
      1               5                  10                  15

Glu Gln Met Ala Ser Lys Lys Arg Glu Leu Phe Ser Asn Leu Gln Glu
                20                  25                  30
```

```
Phe Ala Gly Pro Ser Gly Lys Leu Ser Leu Leu Glu Val Gly Cys Gly
            35                  40                  45

Thr Gly Ala Asn Phe Lys Phe Tyr Pro Pro Gly Cys Arg Val Thr Cys
        50                  55                  60

Ile Asp Pro Asn Pro Asn Phe Glu Lys Phe Leu Ile Lys Ser Ile Ala
    65                  70                  75

Glu Asn Arg His Leu Gln Phe Glu Arg Phe Val Val Ala Ala Gly Glu
80                  85                  90                      95

Asn Met His Gln Val Ala Asp Gly Ser Val Asp Val Val Val Cys Thr
                100                 105                 110

Leu Val Leu Cys Ser Val Lys Asn Gln Glu Arg Ile Leu Arg Glu Val
            115                 120                 125

Cys Arg Val Leu Arg
        130
```

(2) INFORMATION FOR SEQ ID NO: 347:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 49 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -37..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.1
seq LLLYLCCMINIHH/LP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 347:

```
Met Ser Leu Leu His Gly Asn Lys Met Cys Val Thr Ile Arg Pro Thr
        -35                 -30                 -25

Gly Gln Pro Leu Asn Gly Asp Leu Leu Leu Leu Tyr Leu Cys Cys Met
    -20                 -15                 -10

Ile Asn Ile His His Leu Pro Pro Val Val Leu Pro Arg Thr Pro Gln
 -5                  1               5                  10

Gly
```

(2) INFORMATION FOR SEQ ID NO: 348:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -17..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.9
seq ISYFIAFPNLSQA/EL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 348:

```
Met Ser Phe Asn Ile Ser Tyr Phe Ile Ala Phe Pro Asn Leu Ser Gln
    -15                 -10                 -5

Ala Glu Leu Thr His Pro Arg Cys Ser Tyr Thr Gly Leu Ser Ser Ser
     1               5                  10                  15

Cys Gly Phe Gln Leu Ser Asp Thr Pro His Arg
                 20                  25
```

(2) INFORMATION FOR SEQ ID NO: 349:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -20..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq LTIILLPVHLLIT/IY

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 349:

```
Met Lys Leu Lys Xaa Asn Val Leu Thr Ile Ile Leu Leu Pro Val His
-20                 -15                 -10                  -5

Leu Leu Ile Thr Ile Tyr Ser Ala Leu Ile
```

1 5

(2) INFORMATION FOR SEQ ID NO: 350:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 56 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -14..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq AALVTVLFTGVRR/LH

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 350:

```
Met Ala Ala Leu Val Thr Val Leu Phe Thr Gly Val Arg Arg Leu His
                -10                 -5                   1

Cys Ser Ala Xaa Leu Gly Arg Ala Ala Ser Gly Xaa Tyr Ser Arg Asn
         5                   10                  15

Trp Leu Pro Thr Pro Pro Ala Thr Gly Pro Leu Pro Ser Ser Gln Thr
     20                  25                  30

Gly His Met Arg Met Ala Ala Arg
 35                  40
```

(2) INFORMATION FOR SEQ ID NO: 351:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 77 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Surrenals

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -42..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq ILMRDFSPSGIFG/AF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 351:

```
Met Ala Ser Val Gly Glu Cys Pro Ala Pro Val Pro Val Lys Asp Lys
        -40                 -35                 -30

Lys Leu Leu Glu Val Lys Leu Gly Glu Leu Pro Ser Trp Ile Leu Met
    -25                 -20                 -15

Arg Asp Phe Ser Pro Ser Gly Ile Phe Gly Ala Phe Gln Arg Gly Tyr
-10                 -5                  1                   5

Xaa Arg Tyr Tyr Asn Xaa Tyr Ile Asn Val Xaa Lys Gly Ser Ile Ser
            10                  15                  20

Gly Ile Xaa Met Val Leu Ala Cys Tyr Val Leu Phe Ser
        25                  30                  35
```

(2) INFORMATION FOR SEQ ID NO: 352:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Surrenals

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -15..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 10.9
seq ALLVLVTVALASA/HH

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 352:

```
Met Leu Ala Leu Leu Val Leu Val Thr Val Ala Leu Ala Ser Ala His
-15                 -10                 -5                  1

His Gly Gly Glu His Phe Glu Gly Ala
            5                   10
```

(2) INFORMATION FOR SEQ ID NO: 353:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 56 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.9
seq VLLFSGFWGLAMG/AF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 353:

```
Met Arg Ile Ile Ser Arg Gln Ile Val Leu Leu Phe Ser Gly Phe Trp
    -20                 -15                 -10

Gly Leu Ala Met Gly Ala Phe Pro Ser Ser Val Gln Ile Gly Gly Leu
 -5                  1               5                  10

Phe Ile Arg Asn Thr Asp Gln Glu Tyr Thr Ala Phe Arg Leu Ala Ile
            15                  20                  25

Phe Leu His Asn Thr Ser Pro Gly
        30                  35
```

THIS PAGE BLANK (USPTO)